BEST AVAILABLE COPY

Transgenic Research 5, 405–411 (1996)

Expression of the murine wild-type tyrosinase gene in transgenic rabbits

BERNHARD AIGNER[1]‡, URBAN BESENFELDER[1]‡, JANOS SEREGI[2], LASZLO V. FRENYO[2], TIBOR SAHIN-TOTH[2] and GOTTFRIED BREM[1,2]*

[1]*Institut für Tierzucht und Genetik, Veterinärmedizinische Universität Wien, Josef-Baumann-Gasse 1, A-1210 Wien, Austria (Fax: +43-1-25077 5692)*
[2]*Research Institute of the University for Veterinary Science, Üllö, Hungary*

Received 21 July 1995; revised 27 October 1995; accepted 6 November 1995

The tyrosinase gene is known to be essential for melanization and has been shown to rescue pigmentation in albino mice. Previously we have described the strict copy-number-dependent expression of a murine wild-type tyrosinase gene construct over several generations in transgenic mice. In this study, we analysed the same gene construct as a marker gene for the transmission and expression of transgenes in rabbits. Using an albino hybrid strain, we produced transgenic rabbits expressing the murine tyrosinase gene. Strict correlation between integration and expression of the transgene and stable germline transmission of the integrated gene construct according to the Mendelian pattern of inheritance was observed. Thus, breeding control was facilitated by simple phenotypic examination of the transgenic animals. In contrast to mice transgenic for the same gene construct, tyrosinase-transgenic rabbits showed a greater variety in hue, intensity and extent of coat pigmentation, which is caused by the diversity in the loci affecting the melanization. Benefits and limitations of tyrosinase as a marker gene for the detection of homozygous individuals in the albino hybrid strain used are discussed.

Keywords: breeding control; colour marker; gene transfer; pigmentation

Introduction

Pigmentation in mammals is based on the production of melanin, a heteropolymer of different metabolic intermediates of tyrosine. In mice, it is controlled by more than 50 independent loci consisting of more than 150 different alleles (Hearing, 1987). Tyrosinase (EC 1.14.18.1) is the key enzyme in this pathway and is exclusively expressed in melanocytes of the skin and pigment cell layers of the eyes (Silvers, 1979). In mice, the single copy gene maps to chromosome 7 and covers the *c*- (albino-) locus (Hearing and Tsukamoto, 1991). The murine tyrosinase gene has a length of about 70 kb and includes five exons coding for a mRNA of about 2 kb (Ruppert *et al.*, 1988). While the dominant wild-type allele *C* results in full tyrosinase activity, alleles giving rise to reduced melanin production have been described (Halaban *et al.*, 1988; Kwon *et al.*, 1989; Beermann *et al.*, 1990). The lack of enzyme activity is caused by the recessive allele *c* harbouring a defined point mutation in a highly conserved region of the first exon (Shibahara *et al.*, 1990). This point mutation was observed in all albino mouse strains examined (Jackson and Bennent, 1990). The lack of tyrosinase activity masks the information of all other loci affecting the melanization (Hogan *et al.*, 1986). After tyrosinase gene expression in tissue culture systems (Müller *et al.*, 1988; Takeda *et al.*, 1989; Yamamoto *et al.*, 1989) murine wild-type tyrosinase gene constructs have been shown to rescue the albino phenotype in mice (Beermann *et al.*, 1990; Tanaka *et al.*, 1990; Yokoyama *et al.*, 1990).

The molecular genetics of the albino phenotype in rabbits remain to be established. Nevertheless we have demonstrated the rescue of the albino phenotype of ZIKA^R hybrid rabbits by introducing a murine wild-type tyrosinase gene construct (Aigner and Brem, 1993). Having established the possibility for the use of the

‡Present address: Abteilung Biotechnologie in der Tierproduktion, IFA Tulln, Konrad-Lorenz-Str. 20, A-3430 Tulln, Austria.
*To whom correspondence should be addressed.

0962–8819 © 1996 Chapman & Hall

tyrosinase as a marker gene for the transmission of transgenes in co-integration experiments by strict copy-number dependent gene expression in mice (Aigner and Brem, 1994), in this study we test the same murine tyrosinase minigene construct (Beermann *et al.*, 1990) in another species for its use as a marker gene by analysing transmission and expression of the transgene in several transgenic rabbit lines.

Materials and methods

Generation of transgenic rabbits

The albino ZIKA^R hybrid strain was used for the generation of transgenic rabbits. Animal husbandry, embryo production, pronuclear microinjection of 500 copies of the transgene and laparoscopic transfer of the zygotes to the recipients were performed as previously described (Besenfelder and Brem, 1993; Brem, 1993). The gene construct ptrTyr4 has a length of 15.5 kb and includes 5.5 kb of the 5′ untranslated region, exon 1, almost the entire first intron and the exons 2–5 of the murine wild-type tyrosinase gene followed by the SV40 splice and polyadenylation cassette (Beermann *et al.*, 1990).

DNA analysis

Genomic DNA was isolated from tissue probes by standard protocols (Ausubel *et al.*, 1987). Polymerase chain reaction was carried out for the detection of transgenesis (Beermann *et al.*, 1990); transgenic animals showed a specific 515 bp signal. Examination of integration and stable transmission of the transgene by Southern analysis was performed as described (Aigner and Brem, 1994). A 1 kb SV40 fragment (*Bam* HI–*Sal* I) was used as a transgene specific probe (Gorman *et al.*, 1982).

The copy numbers of the transgenic rabbit lines were estimated by slot blot analysis. Dilutions of transgenic DNA (8, 4, 2 and 1 μg genomic DNA) were visually compared to differing amounts of the gene construct (in 4 μg rabbit DNA) representing various copy numbers and ptrTyr4-transgenic mouse probes with defined copy number (Aigner and Brem, 1994, 1995). Genomic DNA of non-transgenic animals served as negative control. The differentiation between homozygous and hemizygous littermates after mating two hemizygous siblings was done in duplicate by semiquantitative PCR as described (Aigner and Brem, 1995) using two different pairs of transgene-specific primers (Beermann *et al.*, 1990). The amount of DNA used in the semiquantitative PCR was evaluated by amplification of a 500 bp fragment deriving from the endogenous rabbit transferrin gene (Banfield *et al.*, 1991) with the primers RATF1 (5′-GCCTTTGTC-AAGCAAGAGACC-3′) and RATF2 (5′-CACAGCAGCT-CATACTGATCC-3′) at 62 °C annealing temperature.

RNA analysis

Total RNA was isolated from tissue probes by standard techniques (Chomcynski and Sacchi, 1987), mRNA by using the 'Quickprep Micro Purification' kit (Pharmacia). Reverse transcriptase (RT) PCR was performed as previously described (Innis *et al.*, 1990; Aigner and Brem, 1994). A 449 bp signal shows the specific expression of the integrated gene construct ptrTyr4.

Results and discussion

Generation and analysis of the transgenic founder rabbits

The albino ZIKA^R hybrid strain was used for the generation of rabbits transgenic for the murine wild-type tyrosinase gene construct ptrTyr4. On average, 15 micro-injected zygotes were transferred per uterus horn, resulting in 3.8 offspring born per recipient. The integration of the gene construct was examined by PCR. Table 1 shows the results of the tyrosinase gene transfer. A total of 27 animals were found having integrated the gene construct. The integration efficiency (transgenics/F_0 examined) was 11.4%, the overall efficiency (transgenics/zygotes transferred; corrected to the number of F_0 animals examined) was 0.66%. Comparable gene transfer efficiencies in rabbits were achieved in other programmes (Brem *et al.*, 1994). All phenotypically examined transgenic rabbit founders expressed the murine gene construct in coat and/or eyes, whereas none of the unpigmented animals harboured detectible parts of the gene construct. In mice, a high percentage of transgene expression was also achieved with different tyrosinase gene constructs (Yokoyama *et al.*, 1990; Overbeek *et al.*, 1991). Thus, we have rescued the albino phenotype in the ZIKA^R hybrid strain by tyrosinase gene transfer.

Table 1. Production of ptrTyr4-transgenic rabbits

Zygotes		*Recipients*		*Newborns*		
microinjected	*transferred*	*total*	*pregnant*	*total*	*examined*[a]	*Transgenic by PCR*
7890	7100 (90%)	229	107 (47%)	411	237	27 (11.4%)

[a]As consequence of health problems and death of newborns independent from the gene transfer programme, only 237 newborns were examined for transgenesis.

Colour, intensity and extent of coat pigmentation of the positive F_0 individuals varied over a wide range, ranging from a small, grey splotch on the back (about 5% of the coat without any visible eye melanization and any detectible transgenesis in skin probes outside of this region) to complete and homogenous, dark coat melanization. It was not possible to determine the exact number of animals showing phenotypic coat mosaicism owing to the appearence of natural, transgene-independent, rabbit coat patterns containing unpigmented coat regions (e.g., 'Dutch Belted'), previously masked by the albino phenotype of the rabbits. These rabbit coat patterns might be caused by modifications of loci controlling the pigmentation in the presence of an active tyrosinase. In pigmented mice, white coat regions were found to harbour no melanocytes or undifferentiated melanocytes in the skin (Takeuchi *et al.*, 1988; Bradl *et al.*, 1991; Mintz and Bradl, 1991).

To establish transgenic lines, four founders were chosen showing complete coat pigmentation (No. 2, 3 and 4) or the 'Dutch Belted' pattern (No. 5) with homogenous melanization of the pigmented coat regions (Fig. 1) and melanization of the eyes. This was based on our observations made in mice with the same gene construct indicating that the extent of coat pigmentation of the founders reflects the germline participation in the integration of the transgene (Aigner and Brem, 1994). A phenotypic coat mosaic (No. 1: map-like extent of the pigmentation over the whole body without sharp borders to the unpigmented regions) was bred as a control, expecting a lower percentage of transmission of the transgene to the offspring. The results agreed with our data derived from the ptrTyr4-transgenic founder mice. The classification according to the Sign test after Dixon and Mood (5%-level) (Sachs, 1992) revealed three of the four completely melanized animals as hemizygous (No. 2, 3 and 5) and the other full pigmented founder (No. 4) harbouring more than one integration site of the transgene, whereas the transgenic founder with the phenotypic coat mosaic (No. 1) showed germline mosaicism (Table 2).

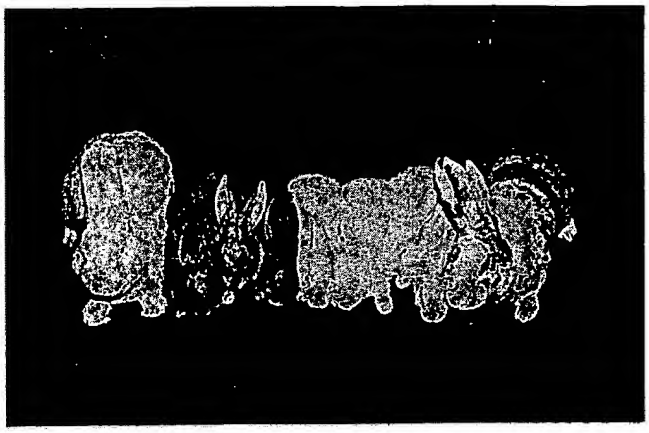


Fig. 1. Founder No. 5 (right, 'Dutch Belted') mated to the non-transgenic ZIKA^R albino rabbit (left) with the F_1 offspring; the three hemizygous F_1 littermates show the same intensity of coat pigmentation.

Expression of the gene construct ptrTyr4

The examination of various tissue probes of an adult hemizygous individual (line 5) showing phenotypically complete and homogenous melanization of coat and pigment cell layers of the eyes by RT-PCR detected the expected expression of the murine gene construct in the pigmented tissues. Transgene expression was also observed in pancreas and testis (Fig. 2). In mice as well, expression of the same gene construct was found in tissues other than those expected (Beermann *et al.*, 1990). This might be caused by the presence of misdirected melanocytes in these tissues, which is found in various species. In addition to these tissues, in other tissue probes (liver and spleen) amplified cDNA derived from mRNA molecules with the small T intron of the SV40 cassette not spliced out was detected. Northern analysis of the same tissue probes showed no detectable bands indicating a low number of transcripts in the positive probes examined. Furthermore, the low number of melanocytes in the skin and different points of examination were suggested to be

Table 2. Production of transgenic lines by breeding transgenic founders to non-transgenic albino rabbits

Founder No.	*phenotype*	*F_1 generation*			*Classification of the founders according to the Sign test*
		total	*transgenic*	%	
1	coat mosaic	25	7	28	germline mosaic
2	completely pigm.	21	10	48	hemizygous
3	completely pigm.	15	7	47	hemizygous
4	completely pigm.	35	26	74	2 independent integration sites
5	'Dutch Belted'	7	3	43	hemizygous
Total: $n = 5$		103	53	–	–

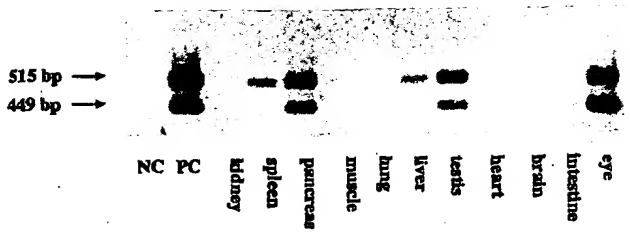


Fig. 2. RT-PCR analysis of various tissue probes of a hemizygous rabbit (line 5). The 1 kb SV40 fragment (*Bam* HI/*Sal* I) was used as a specific probe (Gorman *et al.*, 1982). In addition to the eye, the specific 449 bp signal also appeared in gonads and pancreas. Skin of a non-transgenic mouse was used as negative control (lane NC), pigmented skin of a ptrTyr4-transgenic mouse as positive control (lane PC). Besides these tissues, in liver and spleen amplified cDNA (515 bp) was also detected deriving from mRNA molecules with the 66 bp small T intron of the SV40 cassette not spliced out.

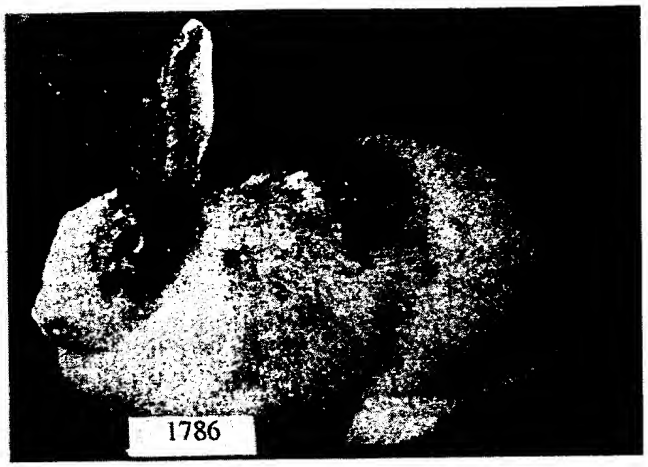


Fig. 3. Hemizygous F_1 rabbit with phenotypic coat mosaicism (line 4A). The phenotype was caused by overlapping of various rabbit coat patterns (such as 'Dutch Belted' and 'Japanese Pattern') indicating a high diversity of loci controlling pigmentation in $ZIKA^R$ hybrid rabbits. In the transgenic rabbits the extent of macroscopically detectible melanization of the pigment cell layers of the eyes varied in a wide range.

responsible for varying results by different groups which have tried to detect the tyrosinase gene expression by northern analysis in the coat of pigmented mice (Müller *et al.*, 1988; Takeuchi *et al.*, 1988; Terao *et al.*, 1989).

Phenotypic examination of different foetal stages showed that tyrosinase-transgenic foetuses can be differentiated from non-transgenic ones already in the mid of gestation (17 days post conceptionem) by the melanization of the pigment cell layers of the eyes (data not shown). Thus, the temporal regulation of tyrosinase during embryonic development, as observed in mice (Beermann *et al.*, 1992), was also rescued in transgenic rabbits.

Correlation between genotype and phenotype in generations F_1 and F_2

A prerequisite for the use of tyrosinase as marker gene is the stable germline transmission and expression of the integrated transgene copies in the following generations. In the F_1 and F_2 generation, a strict correlation between coat pigmentation and detection of the transgene by means of molecular genetic methods was found, i.e. only the pigmented offspring ($n = 95$) inherited the transgene, whereas none of the albino littermates examined ($n = 62$) was detected as transgenic. The transgenic rabbit lines were designated according to the number of the founder animal listed in Table 2. In the F_1 and F_2 generation, transgenic individuals with one integration site of the gene construct appeared to harbour phenotypic coat mosaicism (non-homogeneous melanization with coat regions coloured in different tones and shades). This made it difficult to analyse the correlation between phenotype and genotype of the transgenics (Fig. 3). As observed in mice, the intensity and heterogeneity of coat pigmentation increased with the age of the individuals.

In general ptrTyr4-transgenic F_1 offspring showed a greater variety in hue and extent within the lines compared to the mice transgenic for the same gene construct (Aigner and Brem, 1994). However, the intensity of coat melanization differed only within proportionally confined limits in the lines 1, 2, 3 and 5 harbouring one integration site (phenotypically examined rabbits: line 1: $n = 5$; line 2: $n = 7$; line 3: $n = 4$; line 5: $n = 3$ (Fig. 1)). Line 4 derived from the transgenic founder with two independent integration sites of the gene construct. The transgenic offspring divided by phenotype into four different groups of coat colour intensity (group a–d: $n = 20$). By Southern analysis, the two less intense coloured groups a and b ($n = 11$) were shown to be linked with a reduced signal pattern compared to their darker pigmented siblings of groups c and d ($n = 9$). The reduced copy number of group a and b was confirmed by slot blot analysis and semiquantitative PCR (data not shown). Following the hybridization signal pattern, line 4 split in two lines 4A (with reduced copy number) and 4B. The strict correlation between copy number and expression of the gene construct in the F_1 generation of the transgenic lines was also observed in mice transgenic for the tyrosinase (Aigner and Brem, 1994). Comparison of the transcription levels of the endogenous gene and the transgene was not performed.

The estimation of the copy number by slot blot hybridization revealed about 10 copies of the transgene per cell in hemizygous animals of lines 1, 2 and 3, whereas line 4A harboured approximately 30 copies, line

5 about 60 copies and line 4B 100 copies (data not shown). Different coat colours occured in the differing lines making it difficult to classify the lines by the intensity of coat melanization.

For the production of homozygous animals, two phenotypically similar, hemizygous siblings of the F_1 generation were bred. Fifty-four offspring were born, of which 42 (78%) were pigmented. A strict copy number dependent expression of the tyrosinase gene construct, as seen in mice, was expected to result in two different levels of coat colour intensity: darker pigmented homozygous and lighter melanized hemizygous siblings (Aigner and Brem, 1994). Fig. 4 shows a litter of the F_2 generation, where the phenotypic classification and the differentiation of the transgenic F_2 rabbits in homozygous and hemizygous individuals by semiquantitative PCR matched completely. For other F_2 siblings however, the intensity of coat pigmentation did not always reflect the transgenic status of the animals (data not shown). Moreover, different coat colours and more than two levels of coat colour intensities in the same litter appeared. These deviations of the intensity of coat colour in relation to the genotype within the transgenic lines might not be caused by a copy-number-independent expression of the transgene, but by the ZIKA^R hybrid strain used. After rescuing the pigmentation with the murine wild-type tyrosinase gene construct, the high variety in the phenotype of individuals, as seen in mice, indicated a large diversity in the loci controlling the melanization, which was previously masked in the albino phenotype. The use of a heterologous gene construct seems to play a minor role for the deviations described. Thus, tyrosinase gene transfer in inbred rabbit strains might increase the correspondence between phenotype and genotype of the transgenics by making it easier to relate the intensity of coat pigmentation to the copy number of the animals.

Stability of transmission

In total, 109 transgenic rabbits were examined by Southern analysis for the stability of transmission of the gene construct, including hemizygous rabbits derived from backcrosses of hemizygous F_1 animals to non-transgenic animals ($n = 14$) and the transgenic animals of the F_1 and F_2 generation ($n = 95$). One case of deletion of transgene copies without detectible rearrangements was observed in the F_2 generation of line 5 (data not shown). Compared to the percentage of irregularities of transmission in ptrTyr4-transgenic mice (F_1 and F_2: 0.6%, $n = 509$), the result in rabbits (0.9%) was similar.

Conclusions

In this study we examined tyrosinase as a marker for transgenesis in rabbits by analysing the correlation between phenotype and genotype of tyrosinase-transgenic rabbits. The value of tyrosinase as a marker gene in rabbits was shown by the strict correlation between integration and expression of the transgene and the stable germline transmission of the integrated gene construct according to the Mendelian pattern of inheritance. In addition, a correlation between the intensity of coat pigmentation and the number of integrated transgene copies was shown in the F_1 generation. Although irregularities in coat pigmentation limited the use of tyrosinase for the identification of homozygous individuals in the

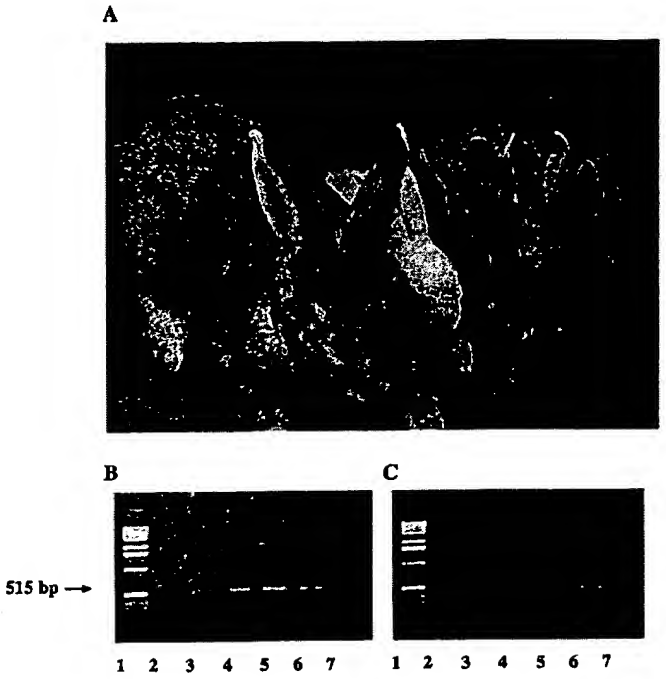


Fig. 4. (A) Generation of homozygous rabbits: F_2 offspring of two hemizygous littermates (line 5) with strict correlation of phenotype and genotype. One hemizygous parent is shown with one non-transgenic albino offspring, two hemizygous (same intensity of coat colour) and one homozygous (darker coloured) offspring. (B) Analysis by semiquantitative PCR revealed a double copy number of the transgene per cell in the darker coloured offspring (lane 5) compared to its hemizygous siblings (lanes 3, 4) and parents (lane 7). As a positive control, double the amount of DNA of the hemizygous parent was used (lane 6), giving rise to a signal intensity comparable to that of the homozygous offspring. DNA of an albino sibling was used as a negative control (lane 2); the molecular weight marker is the 1 kb ladder (Life Technologies, Eggenstein, Germany) (lane 1). (C) The control of the amount of DNA used in the semiquantitative PCR was carried out by semiquantitative PCR for the endogenous rabbit transferrin gene. The application of the probes corresponds to Fig. 4B. As positive control the double amount of DNA of the hemizygous parent was used (lane 6); the negative control was mouse genomic DNA (lane 2). The same signal intensities of the individuals examined show that the same amount of DNA was used in the semiquantitative PCR.

ZIKA^R hybrid strain used, breeding control of transgenic rabbit lines was facilitated by co-transfer and co-integration of the murine wild-type tyrosinase gene construct in accord with previous observations made in transgenic mice (Beermann *et al.*, 1991; Overbeek *et al.*, 1991). Thus, in line 5, co-inheritance of a co-integrated gene construct causing tissue specific production of IGF-1 (insulin-like growth factor 1) in the mammary gland (Zinovieva *et al.*, manuscript in preparation) was observed in 7 generations. The use of rabbits in gene transfer experiments may be required as alternative species to mice for certain applications (Weidle *et al.*, 1991). Furthermore, rabbits combine a short generation interval and low costs for animal husbandry with use as farm animals which prepares them for experiments in genetic engineering, e.g. production of foreign proteins in the mammary gland (Brem *et al.*, 1994).

Acknowledgements

We thank Prof. G. Schütz, Abteilung Molekularbiologie der Zelle I des DKFZ Heidelberg, for kindly providing the gene construct ptrTyr4, U. Huffstadt for the microinjection of the transgene and Dr M. Müller for critically reading the manuscript.

References

Aigner, B. and Brem, G. (1993) Tyrosinase as marker gene and model for screening transgenes in mice and rabbits. *Theriogenology* **39**, 177.

Aigner, B. and Brem, G. (1994) Tyrosinase gene as a marker gene for studying transmission and expression of transgenes in mice. *Transgenics* **1**, 417–29.

Aigner, B. and Brem, G. (1995) Detection of homozygous individuals in gene transfer experiments by semiquantitative PCR. *BioTechniques* **18**, 574–8.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) *Current Protocols in Molecular Biology*. New York: John Wiley and Sons.

Banfield, D.K., Chow, B.K.C., Funk, W.D., Robertson, K.A., Umelas, T.M., Woodworth, R.C. and MacGillivray, R.T.A. (1991) The sequence of rabbit liver transferrin cDNA. *Biochim. Biophys. Acta* **1089**, 262–5.

Beermann, F., Ruppert, S., Hummler, E., Bosch, F.X., Müller, G., Rüther, U. and Schütz, G. (1990) Rescue of the albino phenotype by introduction of a functional tyrosinase gene into mice. *EMBO J.* **9**, 2819–26.

Beermann, F., Ruppert, S., Hummler, E. and Schütz, G. (1991) Tyrosinase as a marker for transgenic mice. *Nucl. Acids Res.* **19**, 958.

Beermann, F., Schmid, E. and Schütz, G. (1992) Expression of the mouse tyrosinase gene during embryonic development: recapitulation of the temporal regulation in transgenic mice. *Proc. Natl Acad. Sci. USA* **89**, 2809–13.

Besenfelder, U. and Brem, G. (1993) Laparoscopic embryo transfer in rabbits. *J. Repr. Fertil.* **99**, 53–6.

Bradl, M., Larue, L. and Mintz, B. (1991) Clonal coat color variation due to a transforming gene expressed in melanocytes of transgenic mice. *Proc. Natl Acad. Sci. USA* **88**, 6447–51.

Brem, G. (1993) Transgenic animals. In Rehm, H.J. and Reed, G., eds., *Biotechnology – Volume 2: Genetic Fundamentals and Genetic Engineering*, pp. 745–832. Weinheim: VCH.

Brem, G., Hartl, P., Besenfelder, U., Wolf, E., Zinovieva, N. and Pfaller, R. (1994) Expression of synthetic cDNA sequences encoding human insulin-like growth factor-1 (IGF-1) in the mammary gland of transgenic rabbits. *Gene* **149**, 351–5.

Chomcynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–9.

Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**, 1044–51.

Halaban, R., Moellmann, G., Tamura, A., Kwon, B.S., Kuklinska, B., Pomerantz, H.P. and Lerner, A.B. (1988) Tyrosinase of murine melanocytes with mutations at the albino locus. *Proc. Natl Acad. Sci. USA* **85**, 7241–5.

Hearing, V.J. (1987) Mammalian monophenol monooxygenase (tyrosinase): purification, properties and reactions catalyzed. In Kaufman, S., ed., *Methods in Enzymology – Metabolism of Aromatic Amino Acids and Amines*, pp. 154–65. New York: Academic Press.

Hearing, V.J. and Tsukamoto, K. (1991) Enzymatic control of pigmentation in mammals. *FASEB J.* **5**, 2902–9.

Hogan, B., Costantini, F. and Lacy, E. (1986) *Manipulating the Mouse Embryo*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (1990) *PCR protocols: a Guide to Methods and Applications*. San Diego, CA: Academic Press Inc.

Jackson, I.J. and Bennett, D.C. (1990) Identification of the albino mutation of mouse tyrosinase by analysis of an in vitro revertant. *Proc. Natl Acad. Sci. USA* **87**, 7010–4.

Kwon, B.S., Halaban, R. and Chintamaneni, C. (1989) Molecular basis of mouse himalayan mutation. *Biochem. Biophys. Res. Commun.* **161**, 252–60.

Mintz, B. and Bradl, M. (1991) Mosaic expression of a tyrosinase fusion gene in albino mice yields a heritable stripped coat color pattern in transgenic homozygotes. *Proc. Natl Acad. Sci. USA* **88**, 9643–7.

Müller, G., Ruppert, S., Schmid, E. and Schütz, G. (1988) Functional analysis of alternatively spliced tyrosinase gene transcripts. *EMBO J.* **7**, 2723–30.

Overbeek, P.A., Aguilar-Cordova, E., Hanten, G., Schaffner, D.L., Patel, P., Lebovitz, R.M. and Lieberman, M.W. (1991) Coinjection strategy for visual identification of transgenic mice. *Transgenic Res.* **1**, 31–7.

Ruppert, S., Müller, G., Kwon, B. and Schütz, G. (1988) Multiple transcripts of the mouse tyrosinase gene are generated by alternative splicing. *EMBO J.* **7**, 2715–22.

Sachs, L. (1992) *Angewandte Statistik*. New York: Springer Verlag.

Shibahara, S., Okinaga, S., Tomita, Y., Takeda, A., Yamamoto, H., Sato, M. and Takeuchi, T. (1990) A point mutation in the tyrosinase gene of BALB/c albino mouse causing the cysteine → serine substitution at position 85. *Eur. J. Biochem.* **189**, 455–61.

Silvers, W.K. (1979) *The Coat Color of Mice – A Model for*

Mammalian Gene Action and Interaction. New York: Springer Verlag.

Takeda, A., Matsunaga, J., Tomita, Y., Tagami, H. and Shibahara, S. (1989) Molecular analysis of the DNA segments cross-hybridizable to the tyrosinase gene in patients afflicted with oculocutaneous albinism. *Tohoku J. Exp. Med.* **159**, 333–40.

Takeuchi, S., Yamamoto, H. and Takeuchi, T. (1988) Expression of tyrosinase gene in amelanotic mutant mice. *Biochem. Biophys. Res. Commun.* **155**, 470–5.

Tanaka, S., Yamamoto, H., Takeuchi, S. and Takeuchi, T. (1990) Melanization in albino mice transformed by introducing cloned mouse tyrosinase gene. *Development* **108**, 223–7.

Terao, M., Tabe, L., Garattini, E., Sartori, D., Studer, M. and Mintz, B. (1989) Isolation and characterization of variant cDNAs encoding mouse tyrosinase. *Biochem. Biophys. Res. Commun.* **159**, 848–53.

Weidle, U.H., Lenz, H. and Brem, G. (1991) Genes encoding a mouse monoclonal antibody are expressed in transgenic mice, rabbits and pigs. *Gene* **98**, 185–91.

Yamamoto, H., Takeuchi, S., Kudo, T., Sato, C. and Takeuchi, T. (1989) Melanin production in cultured albino melanocytes transfected with mouse tyrosinase cDNA. *Jpn J. Gen.* **64**, 121–35.

Yokoyama, T., Silversides, D.W., Waymire, K.G., Kwon, B.S., Takeuchi, T. and Overbeek, P.A. (1990) Conserved cysteine to serine mutation in tyrosinase is responsible for the classical albino mutation in laboratory mice. *Nucl. Acids Res.* **18**, 7293–8.

THIS PAGE BLANK (USPTO)

Thrombosis and Haemostasis © F.K. Schattauer Verlagsgesellschaft mbH (Stuttgart) 78(1) 537–542 (1997)

Current Progress in the Production of Recombinant Human Fibrinogen in the Milk of Transgenic Animals

Stephen P. Butler[1], Kevin van Cott[2], Anuradha Subrumanian, Francis C. Gwazduaskas[1] and William H. Velander[2]

[1]Department of Dairy Science, Virginia Polytechnic Institute and State University, Blacksburg, VA; [2]Department of Chemical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Introduction

Transgenic livestock as an alternative source of recombinant fibrinogen

Many applications for fibrin sealant (FS) have been developed over past two decades due to the increased availability of human fibrinogen (hfib) available from cryoprecipitation processing of human plasma (1). In addition to use as an hemostatic agent, new FS-applications include carrier matrix applications for the delivery of drugs and biologics (1–5). Estimates of annual US clinical need are now greatly in excess of the approximately 300 kg per year which can be harvested using the currently low yield, cryoprecipitation methods on the total of 7 million liters of human plasma available for fractionation in the US (6). Thus, the recombinant source must be capable of economically supplying about 1000 kg/yr in order to replace the maximum possible hfib obtainable from current plasma fractionation by cryoprecipitation. The use of the mammary gland of transgenic livestock as a bioreactor for producing rhfib is currently being evaluated in dairy livestock and pigs. While dairy cattle are the most prodigious milk producers, the goat, sheep, and pig have all demonstrated the ability to produce g/l levels of recombinant proteins. The annual milk yields of these livestock are about 1000 liters per year for the goat (7), 500 l/yr for the sheep, and 100 to 300 l/yr for the pig. Thus, only 2000 sows would be needed to produce 1000 kg/yr or more rhfib at a concentration of 5 g/l in milk. Using the transgenic mammary gland for production of rhfib has several necessary requirements. First, DNA constructs which use mammary specific promoters effective for expressing cDNA, minigene, or genomic fibrinogen coding sequences must be formulated. These constructs are then introduced to embryonic cells, typically using microinjection into the pronucleus of zygotes. Second, the cointegration of constructs containing each of the α, β, and γ encoding sequences for hfib into a transcriptionally responsive chromosomal domain is needed. For the purposes of pharmaceutical production at large-scale, a single cointegration site and associated transgene copy number in that loci is desirable for the facile establishment of a phenotypical and genotypical stable lineage (8). Mosiacism and multiple integration sites frequently occurs in founder animals and this complicates analysis of founder animals (9). Thus, phenotype and genotype can not be reliably defined in transgenic animals until successive generations of offspring obtained from outbreeding with nontransgenic animals are analyzed (8–9).

Regulatory motifs using murine whey acidic protein and ovine β-lactoglobulin promoters to express rhfib

Three different promoters, ovine β-lactoglobulin (oBLG; 10,11) and two versions of the murine whey acidic protein (mWAP; reviewed in 12, W. Velander, unpublished data) have been effective in achieving the co- expression of each rhfib chain at high levels relative to that obtained in cell culture. Both oBLG and mWAP are whey proteins which naturally occur in milk at about 1-2 g/l in the milk of sheep and mice, respectively. However, the levels of recombinant proteins obtained using regulatory elements from the mWAP and oBLG genes have widely ranged from 0.001 to 30 g/l in the milk of different transgenic animal species (reviewed in 12). The structure of the transgene can be used to achieve different levels of expression in the milk of transgenic animals. Hence, the ability to target different levels of expression can be used to optimize the yields of fully functional product relative to limitations encountered in post translational modifications and the secretory pathway (13). To that end, the mWAP promoter has been one of the most effective regulatory elements across species for targeting a relatively narrow range of mammary specific expression range by the use of a specific mWAP-transgene motif (8,13).

Fig.1 shows the structure of oBLG and mWAP hfib-DNA constructs which have been used in efforts to express rhfib in

Correspondence to: W.H. Velander, 142 Randolph Hall, Department of Chemical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0211, USA, Tel: +1-540-2317869; Fax: +1-540-2315022; E-mail: wvelande@vt.edu

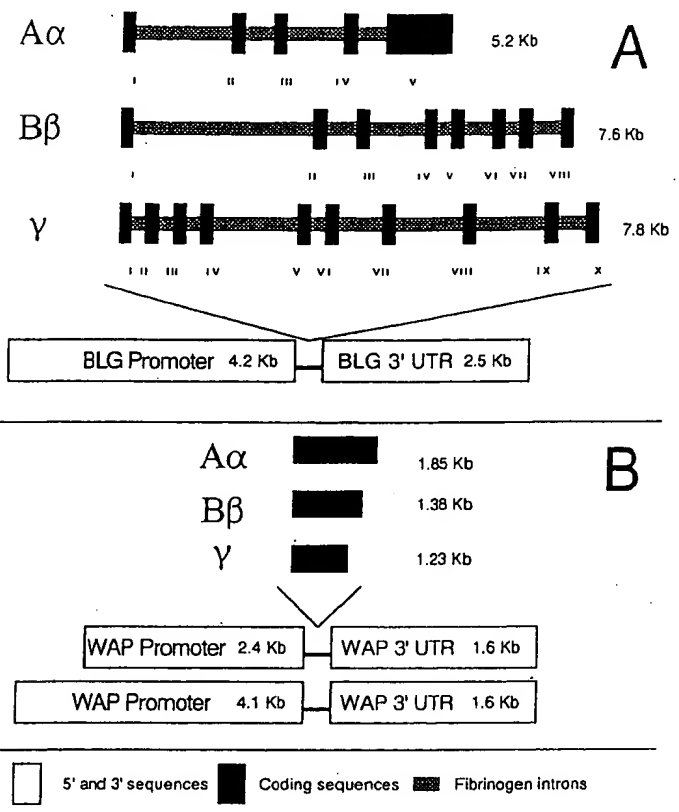


Fig. 1. Examples of human fibrinogen transgene design. A. The genomic regions for each fibrinogen chain (Aα,Bβ,γ) were inserted between 4.2 Kb of the ovine Beta lactoglobulin (BLG) promoter and 2.5 Kb of the BLG associated 3' untranslated region (3'UTR) to produced three individual constructs (adapted from 10). B. The corresponding cDNAs for each fibrinogen chain were inserted between either 2.4 or 4.1 Kb of the murine whey acid protein (WAP) promoter and 1.6 Kb of the WAP associated 3'UTR (Velander , unpublished data).

the milk of transgenic mice, pigs and sheep. A general cloning strategy used in fibrinogen transgene design incorporates inserting a coding sequence of DNA for each individual chain between the milk gene promoter and its associated 3' untransl-ated region (3'UTR). In the examples reviewed here, a total of three separate constructs have been used to express the mature rhfib in the milk. Fig. 1A shows the transgene construct layout for the oBLG driven genomic fibrinogen chains (adapted from 10). The genomic sequences coding for each chain were trimmed of noncoding 5' and 3' flanking sequences and then inserted between the oBLG promoter and 3'UTR DNA sequences. Thus, transcription initiation is regulated by the oBLG promoter with transcription termination and polyad-enylation being controlled by the oBLG-3'UTR. The mWAP transgenes illustrated in fig.1B follow a similar format and have been successful at expressing cDNAs -reviewed in (12).

Expression of rhfib in milk of mouse and pig using different mWAP promoters

The contrast in expression levels obtained in different species has been previously established by results obtained using the 2.4 kb mWAP promoter linked to the cDNA sequence of human protein C (hPC). Transgenic mice having this construct that secreted rhPC in milk at only 0.001 to 0.010 g/l -reviewed in (12). The expression level of rhfib in mouse milk using the same 2.4 kb mWAP format ranged from 0.01 to 0.05 g/l for completely assembled fibrinogen.

One monogenic transgenic mouse line containing only the 2.4 kb mWAP-β-chain-hfib-cDNA construct expressed β- chain rhfib alone at about 0.03 g/l. In contrast to the 2.4 kb mWAP promoter, a longer promoter using 4.1 kb of mWAP and the cDNA of hPC produced recombinant hPC (rhPC) at 0.5 to 1 g/l in the milk of transgenic mice (W. Velander, unpublished data). An expression level of 0.1 to 0.6 g/l rhfib has been obtained in three founder animals having the 4.1 kb mWAP-hfib-cDNA. These observations suggest that the longer mWAP promoter may be able to express cDNAs to relatively high levels in milk of livestock. The two promoters were derived from different genomic mouse libraries and upon sequence analysis, it was determined that a high degree of homology is present in the proximal 1.5 kb with sequence divergence in their distal regions (14). Since both the 2.4 kb and 4.1 kb promoters share a high degree of homology in the proximal region from the transcriptional start, the ability of the longer promoter to express cDNAs at high levels may be due the presence of enhancer elements and/or the lack of repressor elements found in distal region as suggested in (15).

Fig 2A. Shows a western blot analysis of nonreduced SDS PAGE (4 to 12% gradient) analysis of milk from 2.4 kb mWAP-hfib-cDNA transgenic mice. These milks showed a rhfib species of about 340 kDa which was similar to that

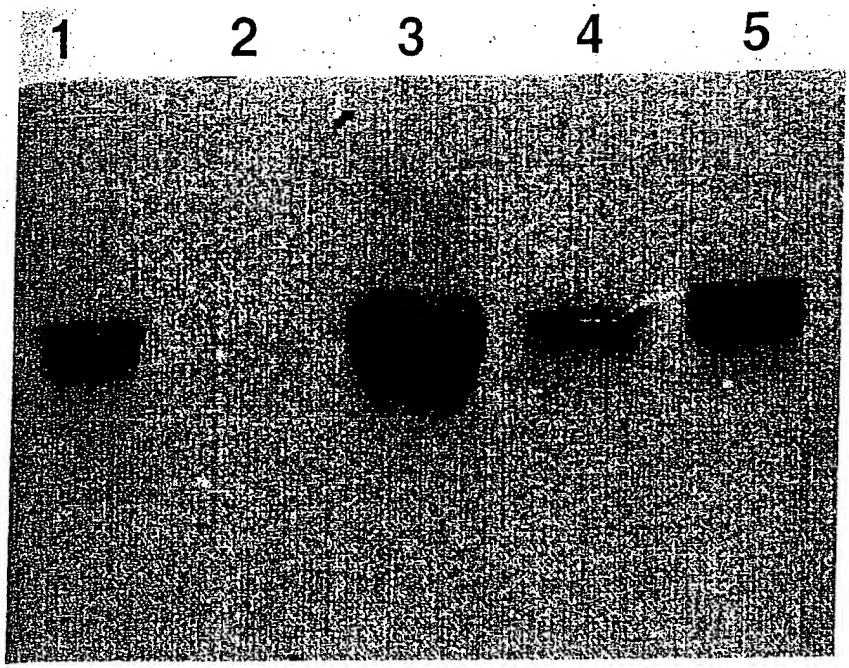


Fig. 2A. Western analysis of partially purified pooled milk samples from 2.4 mWAP-hfib-cDNA, αβγF1 transgenic mice. Transgenic, nontransgenic, and hfib-spiked nontransgenic milk samples were pre-treated by DEAE ion exchange chromatography, followed by $Zn2^+$ precipitation at 8 mM $ZnCl_2$ to selectively remove mouse fibrinogen from the milk (Velander, unpublished data). Mouse fibrinogen partitions into the 8 mM ZnCl2 precipitate, while rhfib remains in the supernatant. Sa-mples were separated on a 4-12% SDS PAGE under non-reducing conditions, transferred to a nitrocellulose membrane, probed with a polyclonal anti-hfib (Celsus, Cincinnati, OH), and visualized by metal enhanced DAB staining (Pierce, Rockford, IL). Lane 1: hfib standard, purified from human plasma, 50 ng. Lane 2: 8mM $ZnCl_2$ supernant from nontransgenic mouse milk, 10 µg total protein loaded. Lane 3: 8mM Zn Cl_2 supernant of nontransgenic mouse milk spiked with 100 ng of hfib. Lane 4: 8mM $ZnCl_2$ supernant of transgenic mouse milk (50 ng rhfib loaded). Lane 5: hfib standard, 100 ng.

obtained for hfib. Fig 2B shows western analyses using a polyclonal antibody detection of SDS-PAGE under reducing conditions for several different F1 offspring from outbreeding of a 4.1 kb-mWAP-hfib-cDNA transgenic founder. These transgenic mice produced about 0.1 to 0.3 g/l rhfib in their milk as detected by polyclonal ELISA. As can be seen in Fig 2B, the expression levels of rhfib are relatively stable throughout lactation. Purified rhfib from the milk of mice containing mWAP constructs was shown to be a functional substrate for dual treatment by both human thrombin and FXIII as clots were obtained and cross-linked fragments appeared in western analysis of the reduced clots (data not shown). Milk from rhfib produced by transgenic mice containing the oBLG-hfib- gene also developed cross-linked fibrin clots after treatment of the milk by human thrombin in the presence of FXIII (10).

Our initial studies in expressing rhfib in the porcine mammary gland have been with cDNA constructs containing the 2.4 kb mWAP promoter. Pronuclear microinjection of equilmolar concentrations of α, β, and γ hfib constructs into one-celled embryos has resulted in 3 litters containing 10 piglets. About 320 microinjected embryos were transfered into a total of 8 synchronized surrogate mothers (40 embryos per recipient). Five of the 10 total piglets born contained at least two of the mWAP-hfib transgenes. Only one of these piglets contained all three transgenes in DNA isolated from tail tissue samples, and it was the only founder animal that contained hfib- transgenes in the germline. Multiple loci were detected in offspring obtained by the outbreeding of this gilt with a nontransgenic boar. Nine of 11 of these offspring contained at least two or three of the α-,β-, and γ-hfib transgenes. Only alpha chain was found in the milk of the founder αβγ-hfib transgenic pig where somatic tissue mosaicism or disfunctional transgene insertion sites may have occurred for some of these transgene insertion sites. Milk from offspring having all three hfib-transgenes will be analyzed to determine if somatic tissue mosiacism or dysfunctional integration sites is occurring in the mammary gland of this single founder animal. Thus, co-integration of separate, multiple transgenes can introduce an increased level of complexity to making and time frame needed for evaluating transgenic livestock.

Expression of rhfib in milk of mouse and sheep using the oBLG promoter

The differences of expression obtained for the 4.2 oBLG promoter using cDNA or gene sequences in mice and sheep are considerable and different than those obtained using mWAP driven constructs in mice and pigs. In contrast to mWAP promoter in mice and pigs, the use of the oBLG promoter to give high levels of expression in the milk of mice and sheep apparently requires genomic coding sequences. Thus, oBLG-hfib-gene constructs were selected in experiments to express rhfib in the milk of transgenic sheep. As also occurred in transgenic mice containing the mWAP-hfib- cDNA constructs, about 85% of oBLG-hfib-gene transgenic mice contained all three hfib-transgenes (10). Transgene copy numbers varied among the different lines but the ratio of the three constructs within a given locus was usually equal. The milk secretion levels observed varied widely from 0.03 to 2 g/l with the average across founding lines being 0.6 g/l. The

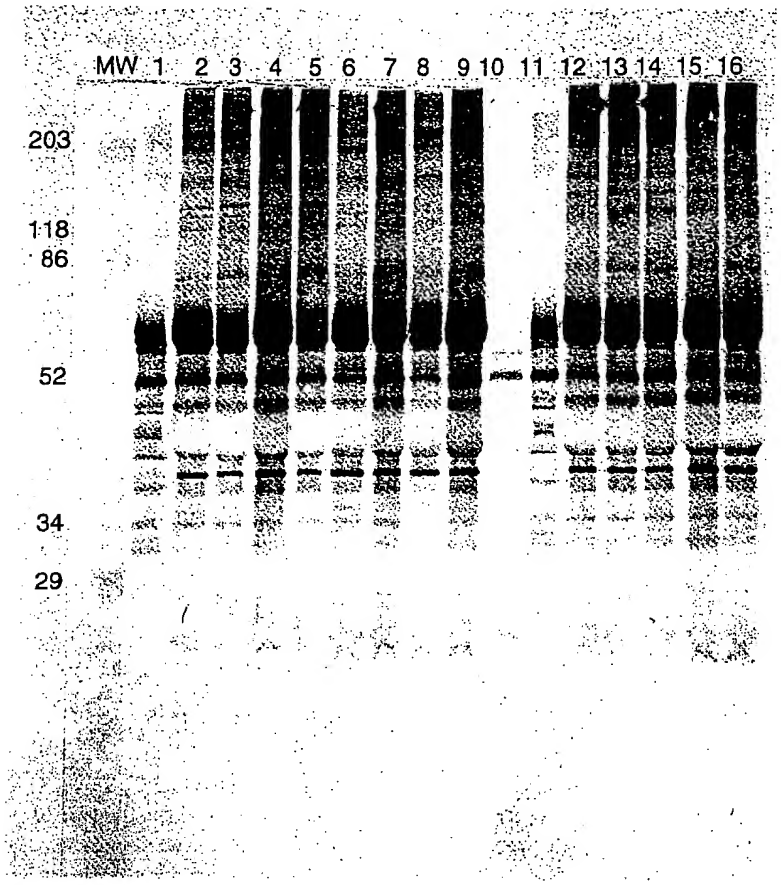


Fig. 2B. Western analysis of daily milk samples from the first (Lanes 2-9) and second (Lanes 12-16) lactations of 4.1 mWAP-hfib-cDNA, αβγ F1 transgenic mice. Equal volumes of defatted milk from transgenic and nontransgenic were applied directly to the gel. Milk samples were electrophoresed under reducing conditions, transferred to a PVDF membrane, probed with a polyclonal antibody to hfib (Celsus, Cincinnati, OH) and visualized with metal enhanced DAB substrate (Pierce, Rockford, IL). Lane 1: hfib standard purified from plasma (100 ng). Lane 2: milk from mouse 10-23, Day 6. Lane 3: milk from mouse 10-23, Day 12. Lane 4: milk from mouse 10-28, Day 7. Lane 5: milk from mouse 10-28, Day 11. Lane 6: milk from mouse 10-13, Day 7. Lane 7: milk from mouse 10-13, Day 11. Lane 8: milk from mouse 10-27, Day 6. Lane 9: milk from mouse 10-27, Day 10. Lane 10: Nontransgenic mouse milk. Lane 11: hfib (100 ng). Lane 12: milk from mouse 10-23, Day 5. Lane 13: milk from mouse 10-28, Day 6. Lane 14: milk from mouse 10-13, Day 5. Lane 15: 10-13, Day 5. Lane 16: 10-27, Day 5.

average expression level for outbred F1 transgenic mice containing the mWAP- hfib-cDNA mice was about 0.3 g/l for 3 lines expressing rhfib at detectable levels. The total number of integrated copies did not correlate with the amount of rhfib secreted, indicating that the constructs are more likely influenced by chromosomal positioning which is commonly observed with transgenes. Pronuclear microinjection of sheep zygotes using the oBLG-hfib genomic constructs resulted in the generation of 9 founding animals (11).

Expression data from four ewes has been reported where 3 animals contained the α-,β-, and γ-transgenes in their genome. A fourth animal contained only genes for the beta and gamma chains. Upon hormone induced lactation at 4 months of age, milk from two of the trigenic animals contained 0.5 g/l and a third ewe contained about 5 g/l. Purified material from the highest expressing ewe was subjected to amino terminal analysis before and after thrombin cleavage resulting in the expected sequences for all three chains. SDS-PAGE analysis under reducing conditions demonstrated that the recombinant fibrinogen can undergo cross-linking to form γ-γ dimers after dual treatment by human thrombin and factor XIII .

Comparisons to rhfib production in cell culture

The production of biologically active hfib requires coordinated expression of three separate fibrinogen genes so that co-translation and assembly of a complex hexameric structure occurs. Transgenic mice containing all three transgenes have secreted a wide range of partially assembled rhfib. For both mWAP-hfib-cDNA and sBLG-hfib-gene transgenes, an estimated range of 10 to 100% of rhfib chains were fully assembled in milks of different transgenic mouse lineages (10, W. Velander. unpublished data). Thus far, no correlation has been established with gene copy number or αβγ transgene ratio and rhfib assembly. Assembly phenomena has been extensively studied in the production of rhfib by mammalian cells in culture (16-19). The mechanism for assembly apparently requires a pool of precursor fibrinogen polypeptides to be retained intracellularly by chaperone proteins which occur within the ER and Golgi secretory apparatus. The cDNAs of human fibrinogen have been expressed in transfected monkey kidney fibroblast (COS1; 17,20), baby hamster kidney (BHK; 19), chinese hamster ovary (CHO; 21) and Hep G2 cells (22). Reports of secretion of individual chains and chain complexes into the media have been limited to individual Aα chain and γ chains, an Aα-γ chain complex and half molecules (αβγ). Our work with monogenic mice containing only α or β cDNAs show that the mammary gland can secrete individual Aα chains and Bβ chains into milk, respectively. In one mosiac, bigenic founder pig containing the 2.4 mWAP α- and β-hFib-cDNA constructs, only the Aα chain polypeptide was secreted. In contrast to

the milk of trangenic mice, individual Bβ chains have not been found as secretion products from any single β-hfib-DNA transfected cell line studied (20). For example, biP chaperones precipitated by anti-biP antibodies from cellular lysates of COS-1 beta-chain-only transfectants contained Bβ chain. This Bβ chain contained a mannose rich, endo-H sensitive carbohydrate suggesting this chain is not transported to the Golgi. Thus, the ability of the mammary gland to secrete separate Bβ chain indicates that its chaperone system operates differently than cell lines which do not secrete Bβ chain alone. The secretory behavior of mammary epithelial cells with rhfib is consistent with the secretion of immature forms of other recombinant proteins into milk such as obtained for pro-rhPC (23). Synthesis rates of functionally active rhfib at 1 pg rhfib/cell/day or less were observed with cultured mammalian cells (18) and about 0.5 pg/cell/day at a cell density of 108 cells/ml using yeast (24). In contrast, density of mammary epithelia is about 10^9 cells/ml while secreting about 0.1 to 1 g/l/hr rhfib in mice and sheep. In summary, one of the advantages of producing rhfib in the milk of transgenic livestock is the 10 and 100 fold higher cell density of the mammary gland relative to yeast and mammalian cells in culture (25).

Summary

The mammary gland of transgenic animals has several advantages for production of heterologous proteins including a high cell density that results in high concentrations of secreted protein. While the mammary gland appears to be able to secrete fully assembled recombinant human fibrinogen (rhfib) at 0.1 to 5 g/l levels, some unassembled rhfib chains are also secreted. Presently, the relationship between unassembled rhfib and the coordinated translation of each nascent rhfib polypeptide in the mammary epithelia is unknown. The secretion of fully and partially assembled rhfib is widely variable among mammalian cell lines and where previously no cell line has been shown to secrete beta chain alone. We have observed that mammary epithelia can secrete Bβ chain into milk as well as immature forms of other recombinant proteins, suggesting it likely uses a different secretory pathway than does the liver. This difference in secretory behavior is possibly due to the natural design of milk, where the precise regulation of post translational modifications and intracellular pools of nascent polypeptides needed to achieve fibrinogen assembly may be less important to fulfill the nutritional function of most milk proteins.

References

1. Alving BM, Weinstein MJ, Finlayson JS, Menitove JE, Frantantoni JC. Fibrin sealant: summary of a conference on characteristics and clinical uses. Transfusion 1995; 35:783-790.
2. Greco F, de Palma L, Spagnolo N, Rossi A, Specchia N, Gigante A. Fibrin-antibiotic mixtures: an in vitro study assessing the possibility of using a biologic carrier for local drug delivery. J Biomed Mat Res 1991; 25:39-51.
3. MacPhee M, Campagna A, Kidd R, Best A, Drohan WN. Fibrin sealant as a delivery vehicle for sustained and controlled release of chemotherapy agents. In: Surgical Adhesives and Sealants: Current Technology and Applications. Sierra D. and Saltz R.,eds. Technomic Publishing Co. Lancaster, PA 1996.
4. MacPhee M, Nunez H, Hennings R, Campagna A, Grummon G, Harding S, Drohan W. Sustained release in vitro and *in vivo* of antibiotics from fibrin sealant: In: Surgical Adhesives and Sealants: Current Technology and Applications. Sierra D. and Saltz R.,eds. Technomic Publishing Co. Lancaster, PA 1996.
5. Lasa C, Hollinger J, Drohan W, MacPhee M. Bone induction by demineralized bone powder and partially purified osteogenin using a fibrin-sealant carrier. In: Surgical Adhesives and Sealants: Current Technology and Applications. Sierra D. and Saltz R.,eds. Technomic Publishing Co. Lancaster, PA 1996.
6. Paleyanda R, Young J, Velander, W *et al.* The expression of therapeutic proteins in transgenic animals. In Hoyer LW and Drohan WN, eds., *Recombinant Technology in Hemostasis and Thrombosis*.New York: Plenum Press. 1991:197-209.
7. Haenlein GFW. Dairy Goat industry of the United States. J dairy Sci 1981; 64: 1288-1304.
8. Van Cott KE, Lubon H, Russell CG, Butler SP, Gwazdauskas FC, Knight J, Drohan WN, Velander WH. Phenotypic and genotypic stability of multiple lines of transgenic pigs expressing recombinant human protein C. Transgenic Res 1997; 6:1-10.
9. Velander WH, Johnson JL, Page RL *et al.* High-level expression of a heterologous protein in the milk of transgenic swine using the cDNA encoding human protein C. Proc Natl Acad Sci USA 1992; 89:12003- 12007.
10. Punkard D, Cottingham I, Garner I, Bruce S, Dalrymple M, Lasser G, Bishop B, Foster D. High-level expression of recombinant human fibrinogen in the milk of transgenic mice. Bio/Technol 1996; 14:867-871.
11. Cottingham I, McKee C, Temperley S, Garner I, Prunkard D, Lasser G, Bishop P, Foster D. Human fibrinogen from the milk of transgenic sheep. In: Tissue Sealants: Current Practice, Future Uses. Cambridge Institute, Newton Upper Falls, MA. March 30 April 2 1996 (abstract).
12. Lubon H, Paleyanda RK, Velander WH *et al.* Blood proteins from tra-nsgenic animal bioreactors.Transfusion Med Rev 1996;10:131-143.
13. Van Cott KE, Williams BL, Velander WH *et al.* Affinity purification of biologically active and inactive forms of recombinant human protein C produced in porcine mammary gland. Journal of Molecular Recognition 1997; 9: 407-414.
14. Paleyanda RK, Zhang D, Hennighausen L, McKnight RA, Lubon H. Regulation of human protein C gene expression by the mouse WAP promoter. Transgenic Res.1994; 3: 335-343.
15. McKnight RA, Spencer M, Dittmer J, Brady JN, Wall RJ, Henninghausen L. An Ets site in the whey acidic protein gene promoter mediates transcriptional activation in the mammary gland of pregnant mice but is dispensable during lactation. Mol. Endocrinol. 1995; 9: 717-724.
16. Yu S, Sher B, Kudryk B, Redman CM. Fibrinogen precursors, order of assembly of fibrinogen chains, J Bio Chem 1984; 259:10574-10581.
17. Hartwig R and Danishefsky KJ. Studies on the assembly and secretion of fibrinogen, J Biol Chem 1991; 266:6578-6585.
18. Roy SN, Procyk R, Kudryk BJ, Redman CM. Assembly and secretion of recombinant human fibrinogen, J Biol Chem 1991; 266:4758-4763.

19. Huang S, Mulvihill ER, Farrell DH, Chung DW, Davie EW. Biosys-thesis of human fibrinogen. J Bio Chem 1993; 268: 8919-8926.
20. Danishefsky K, Hartwig R, Banerjee D, Redman C. Intracellular fate of fibrinogen Bb chain expressed in COS cells. Biochim Biophy Acta, 1990; 1048:201-208.
21. Lord ST, Strickland E, Jayjock E. Strategy for recombinant multichain protein synthesis: fibrinogen Bb-chain variants as thrombin substrates. Biochem 1996; 35:2342-2348.
22. Yu S, Sher B, Kudryk B, Redman CM. Intracellular assembly of human fibrinogen. J. Biol Chem 1983; 258:13407-13410.
23. Lee TK, Drohan WN, Lubon H. Proteolytic processing of human protein C in swine mammary gland. J Biochem 1995; 118:81-87.
24. Roy SN, Kudryk B, Redman CM. Secretion of biologically active recombinant fibrinogen by yeast. J Biol Chem 1995; 270: 23761-23767.
25. Morcol, T, Akers RM, Johnson J.L *et al.* The porcine mammary gland as a bioreactor for complex proteins. Annal NY Acad Sci 1994; 721:218-233.

The

Williar

Past

The use
product
as 1982
from a
physiol
into mo
the pr
integra
transmi
technol
possibl
over a
control
techniq
early e
to isola
of DN
develo
was sh
effects
large
1983,
had b
specifi
The
obviou
pathw
systen
genera
proble
injecti
viabil
transg
cows
Th
severa
prote
tissue
way
gland
milk
prote
(22,2

Corre
Way,
(301)

Transgenic Animals as Bioproducers of Therapeutic Proteins

Juhani Jänne[1,3], Juha-Matti Hyttinen[1,3], Teija Peura[2,3], Minna Tolvanen[2,3], Leena Alhonen[1,3] and Maria Halmekytö[1,3]

Many human therapeutic proteins are currently produced with the aid of recombinant DNA technology in microbial bioreactors and a few also in large-scale animal cell cultures. Although extremely cost-efficient, the microbial production system has many inherent limitations. Micro-organisms, such as bacteria, can read the universal genetic code and hence produce human proteins with correct amino acid sequence, but cannot carry out post-translational modifications, such as glycosylation, or fold the newly synthesized protein properly to ultimately generate a biologically active entity. Moreover, even though the production of the proteins as such is inexpensive, the downstream processing of the final product may be extremely difficult and costly. Many of these disadvantages, especially the lack of post-translational modifications, can be overcome by employing large-scale animal cell cultures for the production of proteins of pharmaceutical interest. However, due to the long generation time and the requirement for rich culture media, the use of animal cell bioreactors is unacceptably expensive. With the advent of transgenic technology, the production of human pharmaceuticals in large transgenic animals has become more and more attractive. The use of targeted gene transfer, the expression of the transgene of interest can be directed to occur in the mammary gland of large farm animals, such as pigs, sheep, goats or dairy cattle, and hence the transgene product is ultimately being secreted into the milk. Although not yet in commercial use, the last few years have witnessed a remarkable progress in this area and proved the feasibility of the use of 'molecular farming' in high-quantity, low-cost production of valuable therapeutic or industrial proteins. While reviewing the progress of the field over the past few years, we discuss in somewhat greater detail aspects connected with the use of dairy cattle as bioproducers of human therapeutic proteins.

Key words: transgenic animals; mouse; sheep; goat; cattle; bioreactor; pharmaceuticals; human erythropoietin; mammary gland; milk.

(Annals of Medicine 24: 273–280, 1992)

Transgenic animals are animals developed from embryos into which foreign genes have been transferred. If the foreign gene is introduced into the one-cell embryo (fertilized oocyte), and if integrated, the transgene becomes a dominant Mendelian genetic characteristic that is inherited by the progeny of the founder animal. The ability to genetically manipulate mammals has opened an immense potential with almost unlimited applications in basic and applied research.

From the [1]Department of Biochemistry & Biotechnology, [2]Department of Anatomy, and [3]A.I. Virtanen Institute, University of Kuopio, Finland.

Address and reprint requests: Dr Juhani Jänne, Department of Biochemistry & Biotechnology, University of Kuopio, P.O. Box 1627, SF-70211 Kuopio, Finland.

The experimental work of the author's laboratory has received financial support from the Ministry of Agriculture and Forestry and from the Academy of Finland.

Production of Transgenic Animals

There are many routes into the germ-line cells (for general reviews see refs 1–3). By far the most widely used is the microinjection of foreign genes into one of the two pronuclei of a fertilized oocyte. Typically, tens to hundreds of gene copies are injected into the pronucleus, of which some will become integrated (probably while repairing chromosomal damage). As a result of integration at the one-cell embryo stage, the foreign

gene potentially occurs in every cell of the born animal. The gene transfer can also be accomplished by retroviral infection of early embryos or transferring the transgene into embryonal stem cells followed by the introduction of the stem cells into blastocysts. All these methods are tedious and require sophisticated equipment (microinjection) or demanding cell culture techniques (embryonal stem cells). Therefore, it is no wonder that profound public interest was aroused by a report in 1989 (4) that foreign genes can be introduced into unfertilized oocytes by sperm cells in connection with *in vitro* fertilization. This observation was, with good reason, named as 'biological cold fusion' as the method would dramatically simplify the production of transgenic animals in general and large animals in particular. Unfortunately, in spite of serious attempts in a number of laboratories all over the world, this finding has so far not been reproduced (5) and hence the comparison with cold fusion appears to be correct.

The first transgenic mice produced by the microinjection technique were generated in 1980 (6) and since then hundreds of transgenic mouse lines have come into existence (1–3). Transgenic mice have established an exciting new experimental approach to mammalian molecular biology. Germ-line transmission of foreign genes makes it possible to study the function and regulation of genes during animal development in their 'natural' environment. More importantly, with the aid of transgenesis it is possible to construct animal models for almost every imaginable human disease ranging from tumourigenesis to a number of metabolic disorders (3). By using fusion gene constructs, a tissue-specific expression directed by various promoters can readily be achieved. Thus, transgenic mice also serve as important models for targeted gene transfer in large domestic animals.

The creation of 'giant' transgenic mice in 1982 (7) harbouring the metallothionein–rat growth hormone fusion gene in their genome not only attracted much public attention, but also led to the realization that transgenic techniques could be applied to the genetic engineering of domestic animals.

Transgenic Livestock

Even though various groups have produced transgenic farm animals, the progress in this field has not been nearly as dramatic as in the generation of transgenic mice. The reasons are obvious: the pregnancy is long, the litter size is small, the availability of fertilized eggs is limited, and technical difficulties are encountered in the microinjection technique. The latter include difficulties in visualizing the pronuclei due to deposition of opaque material in the cytoplasm of oocytes in some species (pigs and cattle). Nevertheless, a number of transgenic farm animals have been generated during the past few years. The giant mouse (7) apparently served as a model for the first transgenic livestock. With the reasonable assumption that insertion of extra copies of growth hormone genes would lead to accelerated growth, the first transgenic pigs harboured either bovine or human growth hormone genes driven by mouse metallothionein (liver-specific) promoter (8). The genes of human growth hormone-releasing factor and human insulin-like growth factor 1 were likewise used (8). Some of these animals indeed grew faster, converted feed more efficiently to body weight, and even had reduced backfat thickness (8). These desirable effects, however, were clouded by a long list of adverse effects: the animals suffered from lameness, lethargy and gastric ulcers; their glucose metabolism was severely disturbed with a striking elevation of plasma insulin (8); the reproductive capacity of pigs expressing the growth hormone transgenes was seriously impaired. As the health problems encountered in these transgenic pigs were believed to be caused by continuous exposure to high circulating levels of growth hormone, more specific gene constructs were used to generate transgenic pigs. A regulatable (by nutritional factors) fusion gene was constructed by fusing the promoter region of rat phosphoenolpyruvate carboxykinase to genomic bovine growth hormone structure gene (9). The resulting transgenic animals showed enhanced feed efficiency and decreased backfat thickness; however, negative characteristics included stress susceptibility, joint pathology and respiratory distress (9).

Transgenic sheep carrying metallothionein–growth hormone gene constructs have likewise been generated (10, 11). Marginal or no growth advantage in these animals was accompanied by serious health problems, such as diabetes and premature death (11).

It thus appears that by using various growth hormone gene constructs, any improvement of the quality of livestock can only be achieved at the cost of severe adverse effects. This is obviously related to the fact that growth hormone exerts an array of metabolic effects unrelated to its growth-promoting activity. It is highly likely that the metabolic effect cannot be entirely avoided even using strictly controllable promoters. Instead of modifying the normal physiology of the whole animal, a targeted, tissue-specific expression of the transgene seems a much more attractive approach. This applies especially to the mammary gland and to the genetic modification of the milk composition in large farm animals.

Genetic Modification of Milk Composition in Large Domestic Animals

Genetic modification of milk composition is based on the use of transgene constructs in which the structural gene of interest is driven by mammary gland-specific regulatory sequences. The milk protein genes appear to contain highly conserved regions, 'milk boxes', in their 5′ flanking regions (12–14). The milk protein genes are expressed exclusively in the mammary gland across the species boundaries. This is exemplified by a recent report (15) showing that the whey acidic protein gene, which only occurs in rodents and rabbits but not in pigs, is effectively expressed in the mammary gland of transgenic pigs. Thus, the expression of any structural transgene under the control of milk protein promoter can be

directed to the mammary gland in all likelihood without affecting the health of the animal.

There are two major approaches to genetic modification of milk composition. It is possible to improve milk quality by introducing more gene copies of milk proteins, to enhance bacterial resistance by transfer of the lysozyme gene, to break down lactose with the aid of lactase gene, or even to change the fat composition by introducing enzymes capable of changing saturated fats to unsaturated (reviewed in refs 14 and 16). The second approach is to create transgenic farm animals secreting proteins of pharmaceutical or industrial value into their milk. The feasibility of the latter approach, i.e. the expression of non-milk proteins under the control of regulatory sequences of a milk protein in the mammary gland, has been demonstrated by a large number of transgenic mouse models. Human tissue plasminogen activator (17), human urokinase (18), a human growth hormone (19) have been produced in the milk of transgenic mice, to mention a few examples. Similarly, human interleukin-2 has been produced in the milk of transgenic rabbits (20). Although the expression level of the human interleukin-2 gene (under the control of rabbit β-casein promoter) was relatively low, the authors suggested that lactating transgenic rabbits may be used as a source of human recombinant proteins, as the transgenic rabbits can be generated in a short time (20). In addition, the rabbit milk (100 g per female per day) has almost three times more protein content than cow's milk (20).

The approach of producing human therapeutic proteins was subsequently extended to include larger farm animals. Several transgenic sheep lines were generated in which human α1-antitrypsin (21) and antihaemophilic factor IX (22) were produced from fusion genes containing sheep β-lactoglobulin gene. Human factor IX is a good example of a human protein that is extensively modified (glycosylation and γ-carboxylation) post-translationally and therefore cannot be produced by bacterial fermentation. The transgenic sheep were produced by using ovine β-lactoglobulin gene with human α1-antitrypsin cDNA or human factor IX cDNA inserted into the 5′ untranslated region of the gene (21, 22). Both cDNAs were expressed, however, very inefficiently. The concentration of factor IX in the milk was 100,000 times lower than the level of endogenous β-lactoglobulin in sheep milk (22). Although α1-antitrypsin was expressed more efficiently (21), the concentrations were apparently much too low to be worth commercial exploitation. The low expression rate of the gene constructs were obviously attributable to the lack of introns, as the introns apparently greatly enhance transcriptional efficiency, at least in transgenic mice (23). As shown later, genomic sequences are expressed strikingly better in transgenic sheep.

In September 1991 three reports were published simultaneously, representing a real breakthrough in the field and demonstrating the feasibility of using farm animals as bioproducers of pharmaceuticals. The Edinburgh group (24) succeeded in generating transgenic sheep producing human α1-antitrypsin in their milk. Unlike their earlier efforts (21, 22), they now used genomic sequences of human α1-antitrypsin operationally linked to the ovine β-lactoglobulin promoter (24). This gene construct worked very well in transgenic mice generating up to 7 g/l of biologically active human α1-antitrypsin in the milk of the mice (25). Two of the transgenic founder sheep secreted human α1-antitrypsin into their milk at the level of 1–5 g/l and the third founder up to 35 g/l (24). In fact, human α1-antitrypsin was the major protein (nearly 50% of total protein) of the milk of the latter animal (24). It is noteworthy that the total milk protein content was almost twice that of normal sheep. The secretion of the recombinant protein remained at this high level throughout the lactation period with no signs of sustained lactation (24). The protein displayed full biological activity and was glycosylated like its plasma-derived counterpart (24). The authors also set up a conventional purification procedure based on ion exchange, dye affinity, hydrophobic interactions and molecular sieving, with the aid of which they achieved more than 95% purification of the protein (24). Finally, the transgenic animals were perfectly normal and healthy (24). This transgenic sheep line is currently in the process of commercialization.

Simultaneously with the report of the transgenic sheep, a U.S. group (26) published their production of transgenic goats harbouring the human tissue plasminogen activator gene (a mutated glycosylation variant) governed by murine whey acidic protein promoter. Two transgenic goats were born representing an integration rate of about 7% (26). The level of plasminogen activator in the milk was only 3 μg/ml, i.e. about 10% of that produced by a recombinant mouse cell line (26). However, the authors reported the birth of a further transgenic goat producing human tissue plasminogen activator under the control of β-casein promoter at levels that were three orders of magnitude higher than those in the first transgenic animals (26). The authors indicated that the level of human tissue plasminogen activator in the milk of the transgenic dairy goat could make it an economically viable bioreactor (26).

The same authors also exploited the feasibility of isolation and purification of the recombinant protein (plasminogen activator) from goat's milk (27). Using conventional purification procedures including acid fractionation, hydrophobic and immunoaffinity chromatography, they obtained, after 8000-fold purification, an apparently homogenous protein displaying 84% of the biological activity of mouse cell-derived plasminogen activator (27). It was calculated that the transgenic animal with the highest level of expression (3 mg/ml) would produce tissue plasminogen activator in 1 day's milk in quantities that are equivalent to a daily harvest of a 1000 l cell culture bioreactor (27). Nevertheless, the authors described the situation as 'a worse case scenario' as regards the expression level and the purification procedure (27).

The third report described the generation of transgenic dairy cattle harbouring the human lactoferrin cDNA in their genome (28). This transgene construct was driven by a 15 kbp-fragment of bovine αS1-casein (the most abundant protein in cow milk) 5′ flanking region and also containing some 3′ flanking region of the casein gene. The construct gave a transgenesis rate of 10% in

mice and about 20% in cattle (28). It is noteworthy that the cattle embryos were produced and fertilized *in vitro* using bovine ovaries obtained from slaughterhouses as the starting material (28). These three different transgenic farm animal species certainly prove the feasibility of this approach and some of them (the transgenic sheep) may even be commercially exploitable.

As we have been working in the production of transgenic dairy cattle, the rest of this article deals with the production of pharmaceuticals in the mammary gland of dairy cattle, with special reference to improvements that could make the whole procedure less labour-intensive and require a smaller number of recipient animals.

Production of Pharmaceuticals in the Bovine Mammary Gland

In Vitro *Maturation and Fertilization of Bovine Oocytes*

Like the Dutch group (28) we routinely collect cattle ovaries from the slaughterhouse, isolate the immature oocytes, mature them *in vitro* (with FSH, LH and oestradiol-17β) and carry out *in vitro* fertilization with frozen thawed bull semen (29, 30). The fertilized bovine oocytes, unlike mouse oocytes, are not transparent, therefore they have to be centrifuged briefly to visualize the pronuclei. The microinjections are carried out with a micromanipulator connected to a microscope equipped with differential interference contrast optics. Figure 1 shows the visualization of one of the pronuclei after displacement of the lipid material (dark) by centrifugation. After the microinjection, the embryos are cultured *in vitro* for a further 6–8 days before subsequent manipulations.

Gene Construct Used

The expression cassette we are currently using is outlined in Fig. 2. The construct is under the control of bovine αS1-casein promoter (about 2.2 kbp) containing the casein signal sequence fused with genomic sequences of the production gene (currently human erythropoietin) and followed by further 3′ flanking sequences of the αS1-casein. αS1-casein is the most abundant protein (10 g/l) (16) in cow milk and its promoter efficiently directs the expression of heterologous genes into the mammary gland of transgenic mice (18). We have generated five transgenic mouse lines harbouring the human erythropoietin gene under the control of bovine αS1-casein promoter. These mice express the construct in the mammary gland and to some extent in the salivary glands (unpublished results). We routinely generate the parts of the gene constructs by amplification of genomic DNA with the aid of polymerase chain reaction. This necessitates the sequencing of each new construct before use as the polymerase chain reaction and the subsequent cloning occasionally generate mutations that have to be corrected.

Sexing and Transgene Integration Analysis of Preimplantation Embryos

At late morula or early blastocyst stage the *in vitro* cultured embryos will be bisected with the aid of a micromanipulator. A biopsy of each embryo is subjected to sexing and transgene integration analysis. The sexing is based on the polymerase chain reaction method we recently adjusted specifically for the purpose of transgenic animals (30) and which so far has proved to be 100% accurate (31). The method to determine whether

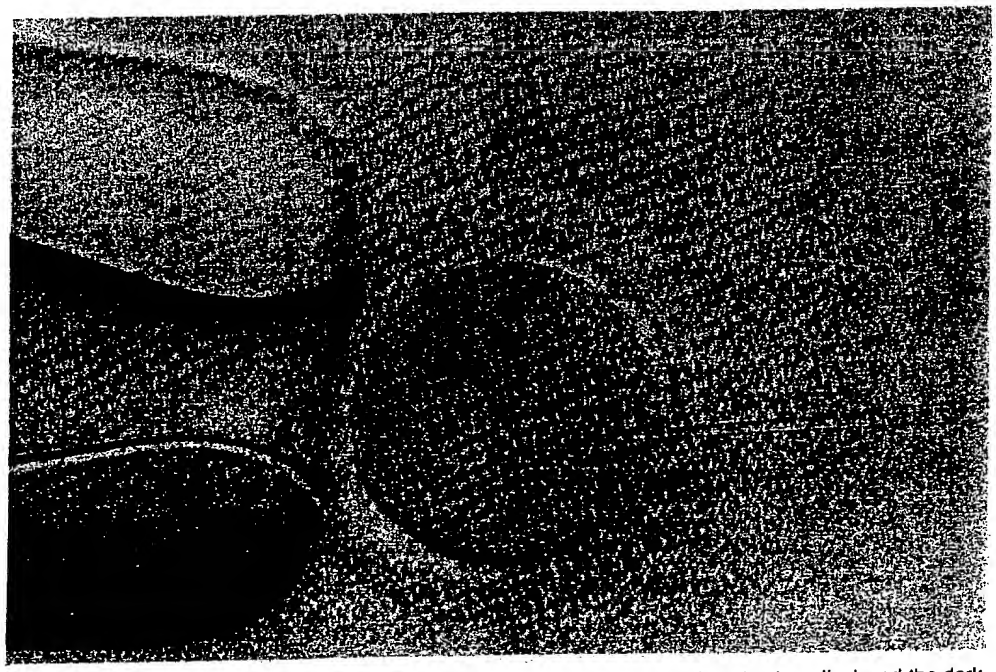


Figure 1. Microinjection of fertilized bovine oocyte. Note that a short centrifugation has displaced the dark cytoplasmic material, allowing visualization of one of the pronuclei.

the transgene has been integrated is based on the use of *Dpn* I restriction endonuclease. This is a unique enzyme as it only cleaves when an adenine belonging to its cleavage sequence (GATC) is methylated (32). The gene constructs are methylated before microinjection with bacterial *dam*-methylase (DNA adenine methylase), which methylates adenine residues at the sequence of GATC. As eukaryotic cells do not possess any maintenance methylase activity for adenine (eukaryotes methylate cytosine), the methylated adenine disappears after the integration and the subsequent replication of the transgene. When the demi-embryo is treated with *Dpn* I, non-integrated transgene is efficiently degraded (contains methyladenine in the cleavage sequence) but integrated and replicated transgene is not. The intact integrated transgene can be shown by using appropriate primers in the polymerase chain reaction. We are currently using a method in which the sexing and integration analysis are combined. A typical combined analysis is shown in Figure 3. Based on the integration analysis, our current integration rate for the human erythropoietin gene construct is about 20% for bovine embryos. After the sex and integration analyses, appropriate demi-embryos will be transferred non-surgically to synchronized recipient animals. The successful verification of the integration of the transgene before the embryo transfer will offer a new dimension for the generation of large transgenic farm animals as the need for recipient animals will be drastically reduced. This, combined with the almost unlimited availability of immature oocytes for maturation *in vitro*, makes the production of transgenic dairy cattle a realistic goal. The whole procedure, i.e. from slaughterhouse to recipient animals, is summarized in Figure 4.

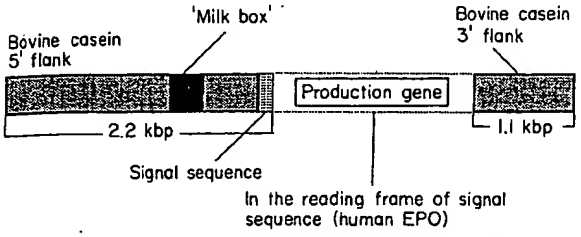


Figure 2. Fusion gene construct in which the structural gene (production gene) is under the control of bovine αS1-casein regulatory sequences. 'Milk box' refers to highly conserved sequences found in the 5' flanking region of milk protein genes in all mammals.

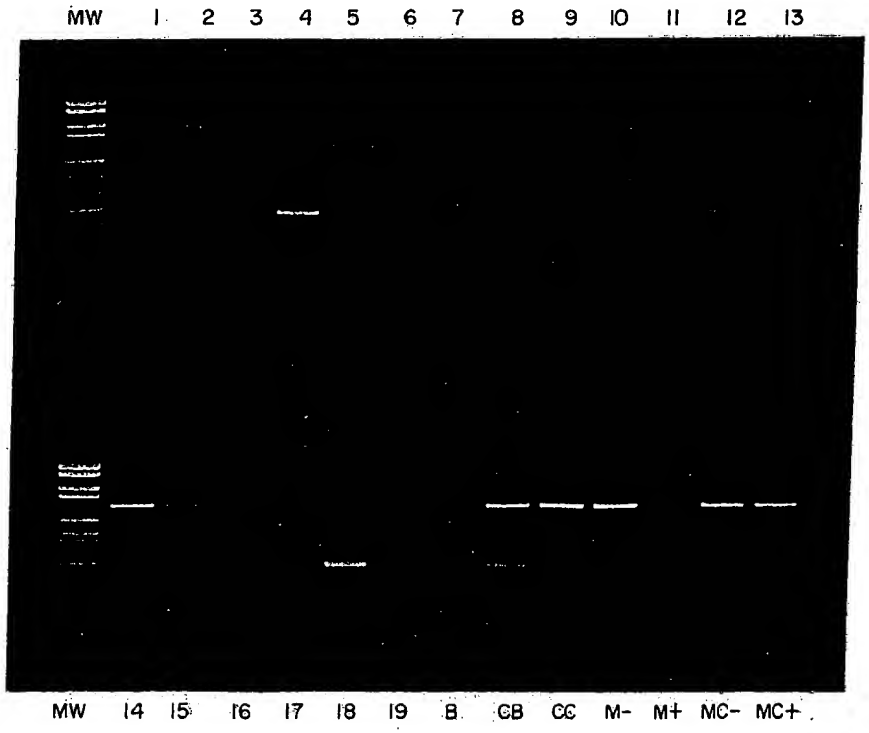


Figure 3. Transgene integration analysis and sex determination of bisected bovine embryos. The lanes are as follows: MW, molecular size markers; B, reagent blank; CB, DNA from transgenic mouse (human erythropoietin) mixed with DNA from bull. The upper fragment represents the transgene integration signal and the lower fragment the male-specific signal; CC, DNA from transgenic mouse (human erythropoietin) mixed with DNA from cow. Note that only the transgene signal is present; M−, microinjected human erythropoietin gene construct with *Dpn* I digestion; M+, microinjected gene construct with *Dpn* I digestion; MC−, DNA, from transgenic mouse mixed with microinjected gene construct without *Dpn* I digestion; MC+, as MC− but with *Dpn* I digestion. Lanes 1–19 represent analyses of bisected bovine embryos. Note that six embryos are positive for the transgene: two (male), three (female), five (male), 14 (female), 15 (male) and 16 (male).

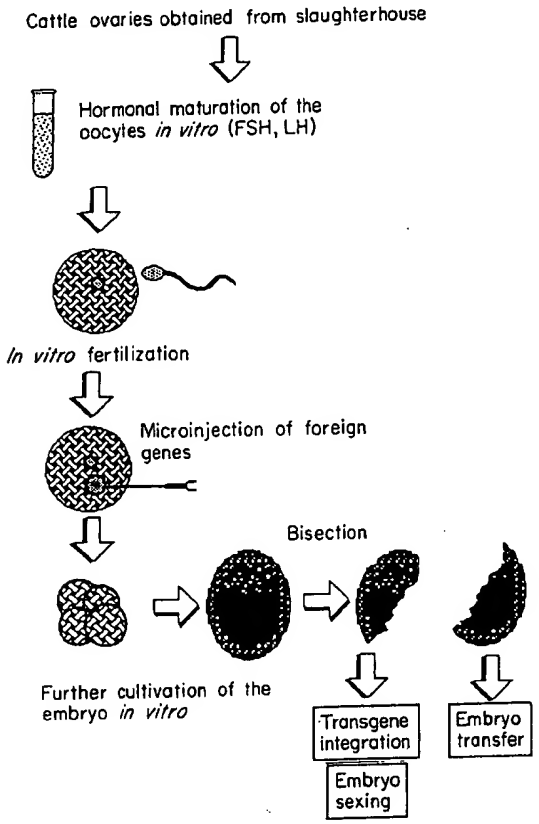


Figure 4. Production of transgenic cattle embryos from oocytes matured and fertilized *in vitro*. Before transfer into recipient animals, the embryos are bisected and one of the demi-embryos is used for transgene integration analysis and sex determination. Only embryos with established transgene integration and preferably of female sex will be transferred.

Possible Ways to Enhance and Control Transgene Expression

As already mentioned, and also experimentally demonstrated in mice (23), genomic sequences are much better transcribed than cDNAs. This probably also applies to large transgenic farm animals, as exemplified by the dramatic increase in the expression of human α1-antitrypsin in transgenic sheep following the replacement of cDNA (22) by genomic sequences (24) in the transgene construct.

The production of transgenic animals by microinjection technique leads to a random integration of the transgene in the embryonal genome and there is generally no direct relationship between the copy number of the integrated transgene and the level of expression (1). This lack of correlation is attributed to a chromosomal position effect that originates from a random integration of the transgene. The entirely unpredictable expression, if any, of a transgene may not be that important in case of transgenic mice, as new lines can be rapidly generated, but it is of utmost importance when large transgenic farm animals with a long pregnancy and small litter size are produced. Position-independent, gene copy number-dependent expression of transgene has been reported in a few cases, such as the human β-globin gene (33), the chicken lysozyme gene (34), the human CD2 gene (35), the human apolipoprotein E gene (36) and the human ornithine decarboxylase gene (37). The position-independent expression of the human β-globin and chicken lysozyme gene has been attributed to so-called matrix attachment elements (locus control regions or dominant control regions) flanking the 5′ and 3′ regions of the transgene and topologically sequestering a functional transcriptional unit (38, 39). It is generally believed that the locus control region is not an enhancer, but rather a new type of regulatory element influencing the organization of chromatin (40). The idea that a transgene construct could be 'shielded' by such regulatory elements has been recently tested using transgenic mice carrying the mammary gland-specific whey acidic protein gene. Interestingly, the inclusion of locus control region (origin not mentioned) in the gene construct led to an improved developmental regulation of the transgene and increased the proportion of the lines that expressed the transgene (41). The possibility of controlling, at least partially, the transgene expression understandably has a profound impact on the generation of large domestic transgenic animals.

Challenges and Opportunities

Although the production of transgenic bioreactors is still in its infancy and we eagerly await commercial applications, the approach certainly is viable. There are, however, many challenges and unsolved problems both concerning the generation of the transgenic bioproducers as well as the downstream processing of the milk to obtain the final product. Key issues in the production of the transgenic animal are the verification of transgene integration prior to the embryo transfer and the assurance of proper transgene integration assay using bisected preimplantation embryos. As indicated earlier, at least a partial solution for the latter problem may be the inclusion of matrix attachment elements in the transgene construct to protect the transgenic locus from a chromosomal position effect. In any event, the production of transgenic farm animals is a major effort both costwise and workwise, and hence it may be advisable to produce several transgene-derived products in the same animal. This can be accomplished by designing expression cassettes that contain two or more independent transcriptional units (structural genes driven by their own heterologous promoters). In fact, we are pursuing this direction by constructing an expression cassette containing transcriptional units for human erythropoietin and for one of the colony-stimulating factors.

As regards the downstream processing, an essential requirement for the transgene product is its stability in the milk. Certain evidence exists that an extensive proteolysis in the milk may not be a problem (24). Difficulties and low yields may also be encountered when purifying the transgene-derived product from the milk. Depending on the rate of expression, large purification factors and modest yields may be the result (27). The physiochemical properties of the milk, i.e. protein-fat

micelles, may likewise create extra problems in the design of proper purification procedures. The fact remains, however, that we are dealing with extremely high volumetric productions (especially dairy cattle) that permit lower purification yields.

The opportunities are almost unlimited. Using proper transgene constructs almost any imaginable protein of pharmaceutical or industrial interest can be produced in transgenic farm animals. Although the therapeutic proteins mentioned in this text are used for the treatment of relatively rare diseases, auxiliary indications for cytokines, for example, are emerging all the time. Thus, human erythropoietin or the colony-stimulating factors are not only used for the treatment of anaemia or leukopaenia but are also used as supportive treatment of cancer, AIDS, etc. (42). Moreover, the number of human therapeutic proteins currently available may look small, but new therapeutic peptides are being discovered continuously. Considering the yearly milk output of dairy cattle (6000–8000 l) and the milk content of αS1-casein (10 g/l), one cow carrying a transgene under the control of αS1-casein promoter would theoretically produce 60–80 kg/yr of the transgene-derived protein. In most cases, these proteins are therapeutically administered in microgram to milligram quantities. Despite having to compromise with modest yields and low expression rates, we would still be within the kilogram business. Add to this the advantages that the bioreactors feed and reproduce by themselves and that the transgenic lines have unlimited potential expansion as the transgene is inherited dominantly in a Mendelian fashion.

We thank Ms. Taru Koponen for her excellent secretarial assistance.

References

1. **Palmiter RD, Brinster RL.** Germ-line transformation of mice. *Ann Rev Genet* 1986; 20: 465–99.
2. **Jaenisch R.** Transgenic animals. *Science* 1988; 240: 1468–74.
3. **Gordon JW.** Transgenic animals. *Int Rev Cytol* 1989; 115: 171–229.
4. **Lavitrano M, Camaioini A, Fazio VM, Dolci S, Forage MG, Spadafori C.** Sperm cells as vectors for introduction of foreign DNA into eggs: genetic transformation of mice. *Cell* 1989; 57: 717–23.
5. **Brinster RL, Sandgren EP, Behringer RR, Palmiter RD.** No simple solution for making transgenic mice. *Cell* 1989; 59: 239–41.
6. **Gordon JW, Scangos GA, Plotkin DJ, Barbosa JA, Ruddle FH.** Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc Natl Acad Sci USA* 1980; 77: 7380–4.
7. **Palmiter RD, Brinster RL, Hammer RE, Trubaner ME, Rosenfeld MG, Birnberg NC, Evans RM.** Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature* 1982; 300: 611–5.
8. **Pursel VG, Hammer RE, Bolt DJ, Palmiter RD, Brinster RL.** Integration, expression and germ-line transmission of growth-related genes in pigs. *J Reprod Fert* 1990; Suppl 41: 77–87.
9. **Wieghart M, Hoover JL, McGrane MM, Hanson RW, Rottman FM, Holzman SH, Wager TE, Pinkert CA.** Production of transgenic pigs harbouring a rat phosphoenolpyruvate carboxykinase-bovine growth hormone fusion gene. *J Reprod Fert* 1990; Suppl 41: 89–96.
10. **Murray JD, Nancarrow CD, Marshall, JT, Hazelton IG, Ward KA.** Production of transgenic Merino sheep by microinjection of ovine metallothionein-ovine growth hormone fusion genes. *Reprod Fertil Dev* 1989; 1: 147–55.
11. **Rexroad CE, Hammer RE, Behringer RR, Palmiter RD, Brinster RL.** Insertion, expression and physiology of growth-regulating genes in ruminants. *J Reprod Fert* 1990; Suppl 41: 119–24.
12. **Hall L, Emery DC, Davies MS, Parker D, Craig RK.** Organization and sequence of the human a-lactalbumin gene. *Biochem J* 1987; 242: 735–42.
13. **Lubon H, Hennighausen L.** Nuclear proteins from lactating mammary gland bind to the promoter of a milk protein gene. *Nucl Acids Res* 1987; 15: 2103–21.
14. **Bremel RD, Yom H-C, Bleck GT.** Alteration of milk composition using molecular genetics. *J Dairy Sci* 1989; 72: 2826–33.
15. **Wall RJ, Pursel VG, Shamay A, McKnight RA, Pittius CW, Hennighausen L.** High-level synthesis of a heterologous milk protein in the mammary glands of transgenic swine. *Proc Natl Acad Sci USA* 1991; 88: 1696–700.
16. **Wilmut I, Archibald AL, Harris S, McClenaghan M, Simons JP, Whitelaw CBA, Clark AJ.** Modification of milk composition. *J Reprod Fert* 1990; Suppl 41: 135–46.
17. **Gordon K, Lee E, Vitale J, Smith AE, Westphal H, Hennighausen L.** Production of human tissue plasminogen activator in transgenic mice. *Biotechnology* 1987; 5: 1183–5.
18. **Meade H, Gates L, Lacy E, Lonberg N.** Bovine alpha-1-casein gene sequences direct high level expression of active human urokinase in mouse milk. *Biotechnology* 1990; 8: 443–6.
19. **Günzburg WH, Salmons B, Zimmermann B, Müller M, Erfle V, Brem G.** A mammary-specific promoter directs expression of growth hormone not only to the mammary gland, but also to Bergman glia cells in transgenic mice. *Mol Endocrinol* 1991; 5: 123–33.
20. **Bühler TA, Bruyère T, Went DF, Stranzinger G, Bürki K.** Rabbit β-casein promoter directs secretion of human interleukin-2 into the milk of transgenic rabbits. *Biotechnology* 1989; 8: 140–3.
21. **Simons JP, Wilmut I, Clark AJ, Archibald AL, Bishop JO, Lathe R.** Gene transfer into sheep. *Biotechnology* 1988; 6: 179–83.
22. **Clark AJ, Bessos H, Bishop JO, Brown P, Haris S, Lathe R, McClenaghan M, Prowse C, Simons JP, Whitelaw CBA, Wilmut I.** Expression of human anti-hemophilic factor IX in the milk of transgenic sheep. *Biotechnology* 1989; 7: 487–92.
23. **Brinster RL, Allen JM, Behringer RR, Gelinas RE, Palmiter RD.** Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci USA* 1988; 85: 836–40.
24. **Wright G, Carver A, Cottom D, Reeves D, Scott A, Simons P, Wilmut I, Garner I, Colman A.** High level of expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. *Biotechnology* 1991; 9: 830–4.
25. **Archibald AL, McClenaghan M, Hornsey V, Simons JP, Clark AJ.** High-level expression of biologically active human α1-antitrypsin in the milk of transgenic mice. *Proc Natl Acad Sci USA* 1990; 87: 5178–82.
26. **Ebert KM, Selgrath JP, DiTullio P, Denman J, Smith TE, Memon MA, Schindler JE, Monastersky GM, Vitale JA, Gordon K.** Transgenic production of a variant of human tissue-type plasminogen activator in goat milk: generation of transgenic goats and analysis of expression. *Biotechnology* 1991; 9: 835–8.
27. **Denman J, Hayes M, O'Day C, Edmunds T, Barlett C, Hirani S, Ebert KM, Gordon K, McPherson JM.** Transgenic expression of a variant of human tissue-type plasminogen activator in goat milk: purification and

characterization of the recombinant enzyme. *Biotechnology* 1991; 9: 839–43.
28. **Krimpenfort P, Rademakers A, Eyestone W, van der Schans A, van den Broek S, Kooiman P, Kootwijk E, Platenburg G, Pieper F, Strijker R, de Boer H.** Generation of transgenic dairy cattle using '*in vitro*' embryo production. *Biotechnology* 1991; 9: 844-7.
29. **Peura T, Aalto J, Rainio V, Vanha-Perttula T.** Pregnancies from bovine oocyte matured and fertilized *in vitro*. *Acta Vet Scand* 1989; 30: 483–5.
30. **Peura T, Hyttinen J-M, Turunen M, Jänne J.** A reliable sex determination for bovine preimplantation embryos using the polymerase chain reaction. *Theriogenology* 1991; 35: 547–55.
31. **Peura T, Hyttinen J-M, Turunen M, Aalto J, Rainio V, Jänne J.** Birth of calves developed from embryos of predetermined sex. *Acta Vet Scand* 1991; 32: 283–6.
32. **Lacks S, Breenberg B.** A deoxyribonuclease of *Diplococcus pneumoniae* specific for methylated DNA. *J Biol Chem* 1975; 250: 4060–6.
33. **Grosveld F, van Assendelft GB, Greaves DR, Kollias G.** Position-independent, high-level expression of the human β-globin gene in transgenic mice. *Cell* 1987; 51: 975–85.
34. **Bonifer C, Vidal M, Grosveld F, Sippel AE.** Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. *EMBO J* 1990; 9: 2843–8.
35. **Greaves DR, Wilson FD, Lang G, Kioussis D.** Human CD2 3′ flanking sequences confer high-level, T cell-specific, position-independent gene expression in transgenic mice. *Cell* 1989; 56: 979–86.
36. **Simonet WS, Bucay N, Lauer SJ, Wirak DO, Stevens ME, Weisgraber KH, Pitas RE, Taylor JM.** In the absence of a downstream element, the apolipoprotein E gene is expressed at high levels in kidneys of transgenic mice. *J. Biol Chem* 1990; 265: 10809–12.
37. **Halmekytö M, Alhonen L, Wahlfors J, Sinervirta R, Jänne OA, Jänne J.** Position-independent, aberrant expression of the human ornithine decarboxylase gene in transgenic mice. *Biochem Biophys Res Commun* 1991; 180: 262–7.
38. **Phi-Van L. Strätling WH.** The matrix attachment regions of the chicken lysozyme gene co-map with boundaries of the chromatin domain. *EMBO J* 1988; 7: 655–64.
39. **Stief A, Winer DM, Strätling WH, Sippel AE.** A nuclear DNA attachment element mediates elevated and position-independent gene activity. *Nature* 1989; 341: 343–5.
40. **Li Q, Zhou B, Powers P, Envert T, Stamatoyannopoulos G.** Primary structure of the goat β-globin locus control region. *Genomics* 1991; 9: 488–99.
41. **Wall RJ, McKnight RA, Shamay A, Hennighausen L.** Matrix attachment sequences improve genetic control of a mammary gland transgene in mice. *Theriogenology* 1992; 37: 319 (abstr).
42. **Niskanen E.** Hematopoetic growth factors in clinical hematology. *Ann Med* 1991; 23: 615–24.

Appl Microbiol Biotechnol (1996) 46:1–9

© Springer-Verlag 1996

MINI-REVIEW

G. Sawers · M. Jarsch

Alternative regulation principles for the production of recombinant proteins in *Escherichia coli*

Received: 17 November 1995/Received revision: 9 February 1996/Accepted: 4 March 1996

Abstract Established expression vectors exploiting regulated promoters such as the *lac* or *tac* promoters have economic and technical limitations when used for the industrial production of recombinant proteins. Consequently, alternative expression systems are being developed that can be more readily manipulated while maintaining high yields of protein. Several suitable expression vectors have been described for use in *Escherichia coli* that are based on promoters the activity of which is under metabolic control. This article discusses the advantages and disadvantages of a cross-section of these expression systems, how they compare with established systems and how they can be applied to the industrial-scale production of recombinant proteins.

Introduction

A number of criteria must be considered when optimizing conditions for the large-scale overproduction of a recombinant protein. These cover the stability of the mRNA (Gross 1989), the efficiency of mRNA translation (Gold and Stormo 1990), the accuracy of amino acid incorporation (Santos and Tuite 1993), whether the protein is correctly folded (Hockney 1994), the formation of insoluble protein aggregates (inclusion bodies) (Hockney 1994), the susceptibility of the product to proteolysis (Gottesman 1990; Nygren et al. 1994) or the requirement for post-translational modification such as proteolytic processing or phosphorylation, and whether the product must be exported to produce an active protein (Hockney 1994; Missiakis et al. 1993). All of these criteria must be considered for each product individually. The choice of promoter used in a vector system has a major bearing on many, but not all, of these criteria. Hence, factors such as inclusion-body formation, the frequency of mistranslation events or the correct folding of a protein can be influenced according to how the activity of the promoter is regulated. Furthermore, use of a strong promoter that produces large amounts of mRNA substrate can compensate in part for mRNA instability, poor translation efficiency or an unstable product. Prudent use of a promoter, therefore, can have a fundamental impact on the quality and yield of a recombinant protein while concomitantly minimizing both development and production costs. In this review we shall describe several metabolically regulated promoters that have been successfully used in expression vectors for the large-scale production of recombinant proteins in *E. coli*. The activity of these promoters responds to fluctuations in metabolism brought about by manipulating cultivation parameters such as the concentration of phosphate or oxygen. Consequently, expression vectors that include metabolically regulated promoters provide cost-effective alternatives to established expression systems based on, for example, the *lac* or *tac* promoters.

G. Sawers
Nitrogen Fixation Laboratory, John Innes Centre,
Norwich NR4 7UH, UK

M. Jarsch (✉)
Div. Therapeutika, F & E Biotechnologie,
Boehringer Mannheim GmbH, Werk Penzberg,
Nonnenwald 2, D-82372 Penzberg, Germany.
Fax: (49) 8856 602144

Categories of recombinant protein

In essence, there are three major categories of recombinant proteins currently produced on an industrial scale, and these are listed in Table 1. The first category includes technical enzymes and proteins for food processing where it is not critical that the product is absolutely pure but high yield is an important factor (Richter 1986). Therefore, since minimal downstream processing is advantageous it is preferable that the

Table 1 Principle categories of recombinant protein

Category	Qualitative and Quantitative prerequisites	Application
1	High yield Low cost Purity is not a major concern	Technical enzymes, e.g. proteases and lipases for washing powders Proteins for food processing or supplementation, e.g. glucose oxidase Biocatalysts, e.g. glucose oxidase
2	High purity Low development effort and cost Yield less critical than for category 1	Enzymes and proteins for in vivo diagnostics, e.g. cholesterol oxidase, glucose dehydrogenase or penicillin-G acylase
3	Very high quality Consistency in production Process must be compatible with Food and Drug Administration guidelines	Human therapeutics, e.g. recombinant tissue plasminogen activator, insulin

protein is produced in high amounts, of the order of 20%–40% of the total cellular protein. For this purpose it is necessary to have an expression vector with a strong promoter. In the other two categories there is a stronger emphasis on the quality of the product, quantity being a secondary consideration (Kopetzki et al. 1994). This is particularly relevant for production of therapeutics where a consistently high-quality product is a prerequisite. Again, this can be affected by the promoter driving expression, particularly, how strong the promoter is and how its activity is controlled. The protein-synthetic capacity of a cell limits the amount, and to a certain extent the quality, of a recombinant protein that can be synthesized. Therefore, although under ideal circumstances it would be desirable to have a host cell that grows rapidly and attains very high cell densities before synthesis of the recombinant protein is switched on, this is not always practicable because the host cell must deliver sufficient ATP and metabolic intermediates to synthesize large amounts of a recombinant protein. Moreover, at very high cell densities nutrients become limiting, which can lead to misincorporation of amino acids (Santos and Tuite 1993) or premature termination of polypeptide chain elongation (Balbas and Bolivar 1990). Hence, it is necessary to be able to control promoter activity easily and efficiently so that synthesis of the recombinant protein is optimal.

Relevant features of promoters in expression vectors

There are both essential and desirable features of a promoter for use in an expression vector (Table 2). It is essential when the promoter is activated that formation of the open transcription complex is efficient and that promoter clearance is rapid. This ensures that large amounts of mRNA are synthesized. The promoter should have sufficient strength to ensure that, under optimal conditions, the ultimate product can attain levels greater than 10% of the total cellular protein, and the promoter should be regulated. The form this regulation takes, the "tightness" (i.e. the extent to which promoter activity can be prevented) of that regulation and the extent to which promoter activity can be induced (the induction ratio) are important considerations.

Transcription initiation from a promoter can be regulated either positively or negatively (Fig. 1). Positive regulation means that a specific activator protein must be present either to permit RNA polymerase to initiate transcription or to increase the frequency of transcription initiation (Gralla 1990). The activity of

Table 2 Features of a promoter desirable in an expression vector

Salient features of the promoter	Alternatives, variables or requirements
Location and stability	1. High-copy-number plasmid 2. Transcription should not interfere with plasmid replication
Strength	1. Slow initiation and rapid elongation, e.g. *lacUV5* promoter 2. Fast initiation and very rapid elongation, e.g. phage T7 promoters
Regulation	1. No promoter activity until product is desired 2. Low-level expression from promoter during growth followed by controlled activation
Activation	1. Removal of a repressor 2. Activation of a positive control factor (activator)
Mode of controlling activation	1. Temperature-shift inactivation of a repressor molecule 2. Chemical inducer added to the culture medium 3. Nutrient deprivation allowing derepression or activation of the promoter

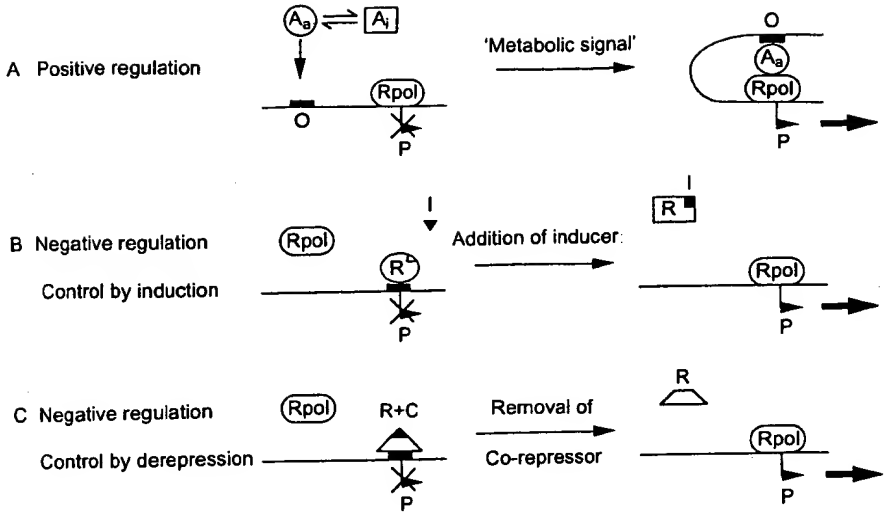


Fig. 1A–C Schematic representation of common modes of promoter regulation. **A** Positive regulation. RNA polymerase (*RPol*) can interact with the promoter (*P*) but cannot form an "open" complex (melt duplex DNA and initiate transcription) because the activator A_i is in the inactive form. Upon responding to a metabolic signal, e.g. a change in oxygen, nitrogen or phosphorus concentration, A_i undergoes a conformational change to its active conformation, A_a. A_a can then bind to a specific operator sequence (*O*; ■), which is located upstream of the promoter on the DNA. By making contact with RNA polymerase facilitated through DNA "looping" the activator can promote open complex formation and transcription initiation. Two modes of negative regulation are depicted. **B** In control by induction, the interaction of RNA polymerase with the promoter is prevented by binding of the repressor protein (*R*) to its operator site, which in this case overlaps the promoter. The repressor has a much higher affinity for its operator than RNA polymerase has for the promoter. The repressor-inducer (I) complex has a drastically reduced affinity for the operator allowing access of RNA polymerase to the promoter. **C** In the case of control by derepression, the scenario is similar to that depicted in **B** except that the repressor must interact with a small molecule (*C* = co-repressor) to enhance its affinity for its operator. Metabolism of the co-repressor inactivates the repressor and liberates the promoter.

the activator protein is usually controlled in response either to a change in the metabolic status of the cell or to the addition of a specific inducer molecule. Negative regulation means that transcription initiation from a promoter is prevented by a repressor molecule (usually a protein). The activity of a negatively regulated promoter, in turn, can be controlled in two ways. In one instance the promoter is controlled by induction, e.g. the *lac* promoter (Makoff and Oxer 1991), where the addition of the inducer lactose prevents repressor binding, thus allowing transcription to proceed. In contrast, the *trp* promoter (Squires et al. 1975), for example, is controlled by repression. Here tryptophan is a co-repressor and, when it becomes depleted, the repressor can no longer bind to the promoter and transcription can occur. How tightly a promoter is regulated depends principally to what extent transcription still occurs in the presence of a repressor or in the absence of an activator. Hence, if the recombinant protein is toxic to the host it is necessary that the promoter is very tightly regulated (Wülfing and Plückthun 1993). In the case of plasmid-based promoters controlled by repression this may require that extra copies of the gene encoding the repressor are also supplied on a plasmid to prevent repressor titration (Stark 1987).

There are several other features of a promoter that may be desirable but this will depend very much on the quality of the product, the quantity of the product that is required and whether the product is toxic to the host cell. These features are listed in Table 2.

Expression systems

Established expression systems

Strong, regulated promoters commonly employed in both research laboratories and industry to drive heterologous gene expression include the promoters from the *lac* operon and the tryptophan (*trp*) biosynthetic operon, as well as phage promoters such as the $\lambda\, p_L$ promoter and the $\phi 10$ promoter from phage T7 (Table 3). All of them have been used to produce large numbers of recombinant proteins that attain levels of at least 5%–10% of the TCP. The *trp* and *lac* promoters are both negatively regulated and have been used successfully to overproduce a vast number of recombinant proteins (Balbas and Bolivar 1990; Tacon et al. 1980; Yansura and Henner 1990).

Hybrid promoters, combining different portions of the *lac* and *trp* promoters have been constructed and used to design improved expression vectors compared with those based around the natural promoters. Examples include the *tac*, *trc* and *tic* promoters (DeBoer et al. 1983; Brosius et al. 1985) which combine the −35 RNA

Table 3 Established expression systems. *ITPG* isopropylthiogalactoside, *IAA* 3-indoleacrylic acid

Promoter	Regulation principle	Induction ratio	Advantages	Disadvantages	References
p_{trp}	Repressed by TrpR-tryptophan complex Starvation of Trp relieves repression IAA mimics Trp starvation	50- to 80-fold	Strong promoter	IAA is expensive Defined synthetic medium required if IAA is not used Potential problems with mistranslation	Squires et al. 1975; Tacon et al. 1980
p_{tac}	Repressed by LacI Induction by lactose or the gratuitous inducer IPTG Catabolite regulation	1000-fold	Strong promoter Very tight regulation IPTG is not metabolized Rapid promoter activation	IPTG is expensive Conditions must be optimized for each scale-up step in the fermentation Inclusion body formation	Gralla 1990; Squires et al. 1975; Fuller 1982
λp_L	Temperature-sensitive cI repressor	> 300-fold	Strong promoter Tight regulation	High energy expenditure Difficult to control Inclusion-body formation	Remaut et al. 1981; Rosenberg et al. 1983
Phage T7 promoters, e.g. $\phi 10$	Transcribed exclusively by phage T7 polymerase	> 1000-fold	Very strong promoter Exclusive transcription	Phage T7 RNA polymerase must be introduced into host cell and its synthesis controlled Owing to extremely high mRNA production, problems with aberrant translation can occur	Studier and Moffat 1986; Studier et al. 1990; Tabor and Richardson 1985

polymerase recognition region from the *trp* promoter with a canonical −10 RNA polymerase recognition sequence and the LacI operator from the *lac* promoter. The three promoters differ in the spacing between the −35 and −10 sequences which affects promoter strength. The *tac* promoter is the most efficient of the three, being five times stronger than *lacUV5* and it still retains the regulation by LacI (DeBoer et al. 1983; Brosius et al. 1985).

Although these expression systems are used in industry there are several shortcomings that detract from their positive attributes. As mentioned above, temperature shifts may cause localized overheating within the fermenter and they are difficult to control. Such shifts have the added disadvantage that correct protein folding is often impaired and they can exacerbate the risk of proteolysis (Hockney 1994). Inclusion-body formation also can be increased by temperature shifts; however, this is not always undesirable and depends very much on the product. Owing to the increased number of operator sites in plasmid-based expression systems it is often necessary to increase the number of repressor molecules. This can be achieved by introducing the gene encoding the repressor onto the expression plasmid itself or by using a second plasmid. It is possible that either of these solutions may reduce host cell growth rates and yields or it might decrease the genetic stability of the expression system (Balbas and Bolivar 1990). Addition of inducing agents, such as indoleacrylic acid or isopropylthiogalactoside is expensive and they must be distributed quickly and evenly throughout the fermenter. Also, care must be taken to ensure that they are completely removed from the product at the processing stage, since their concentration can be relatively high, they are not metabolized and it is possible that they could co-purify with recombinant proteins.

Alternative expression systems

Alternative expression vectors that include promoters the control of which may be more tractable to the cost-effective production of recombinant proteins on an industrial scale have been developed in recent years. The promoter should fulfil the criteria listed in Table 2. Ideally, activation should occur after minor adjustments have been made to the fermentation conditions, for example by altering the flow of oxygen to the culture or by taking advantage of nutrient limitations, such as carbon, phosphorus or nitrogen source depletion, which occur during the fermentation process. The next sections describe several promoters exhibiting metabolic control that have been examined for their capacity to provide possible alternatives to established systems. The characteristics of these promoters are summarized in Table 4 and their mode of regulation is shown in Fig. 2. In some cases the efficacy of the

Table 4 Expression systems under metabolic control

Promoter	Induction principle	Induction ratio	Yield (%) of total cellular protein	References
phoA alkaline phosphatase	Phosphate deprivation	> 1000-fold	20–60	Wanner 1993; Carter et al. 1992
ugp sn-glycerol-3-phosphate transport operon	Phosphate deprivation	~ 100-fold	50	Wanner 1993; Su et al. 1990, 1991; Kasahara et al. 1991; Jarsch unpublished
araB arabinose operon	Glucose depletion and arabinose addition	1200-fold	15	Cagnon et al. 1991; Lobell and Schleif 1991
mgl methyl galactoside transport operon	Glucose addition	~ 100-fold	> 50	Schumacher et al. 1988; Müller 1989; Death and Ferenci 1994; Jarsch, unpublished
vhb Vitreoscilla haemoglobin gene	Microaerobiosis	30-fold	15	Khosla and Bailey 1988; Dikshit et al. 1990
nirB nitrite reductase operon	Anaerobiosis	100-fold	> 30	Charles et al. 1992; Chatfield et al. 1992; Schroeckh et al. 1992
pfl pyruvate formate-lyase operon	Anaerobiosis	25- to 30-fold	> 40	Sawers 1993; Oxer et al. 1991

promoter for use in an expression vector has been tested for only one or two products. However, this suffices to give an overall impression of the qualities and limitations of each and their potential for future development.

Promoters

Nutrient-regulated promoters

The gene encoding alkaline phosphatase (*phoA*) is expressed in *E. coli* at very high levels when cells are starved of inorganic phosphate (Wanner 1993). Expression can be induced more than 1000-fold with PhoA attaining levels of up to 6% of the total cellular protein from a chromosomal copy of the *phoA* gene. The *phoA* promoter is regulated both positively and negatively by the PhoB protein in response to alterations in the inorganic phosphate concentration (Table 4, Fig. 2) (Wanner et al. 1988). PhoB is the regulator component of a two-component signal-transduction cascade and, when the inorganic phosphate concentration [P_i] is above 5 mM, PhoB binds to the promoter and prevents transcription. If [P_i] drops below 1 mM PhoB becomes phosphorylated by the PhoR histidine kinase, which converts PhoB protein into a transcriptional activator. Specific details of Pho regulon control have been reviewed recently (Wanner 1993).

The *phoA* promoter has been used to construct phosphate-regulated expression vectors that have been successfully employed to produce recombinant proteins. Two examples include the production of human epidermal growth factor (Oka et al. 1985) and humanized Fab' fragments (Carter et al. 1992). In the latter study the concentration of Fab' secreted into the culture medium attained levels of 1–2 gl^{-1}.

The promoter from the *ugpBAECQ* operon, encoding an *sn*-glycerol-3-phosphate transport system, is also activated by PhoB phosphate in response to phosphate starvation (Wanner 1993; Su et al. 1990). Maximal promoter activation is also dependent on cAMP receptor protein (CRP) (Kasahara et al. 1991; Su et al. 1991). Reduction in both the phosphate and glucose concentration is therefore required to activate *ugp* promoter transcription. A pBR322-based expression vector has been developed (Su et al. 1990) which is applicable to large-scale fermentations (Table 4). The *ugp* promoter is very strong and exhibits both equivalent strength, under inducing conditions, and induction ratios to those of the *tac* promoter.

Phosphate-regulated promoters are essentially silent at high phosphate concentrations and are thus useful for expression of genes with products that may be toxic to *E. coli* at high concentration. Response to phosphate starvation is rapid, with maximal induction occurring within 30–60 min after the onset of nutrient limitation. The phosphate concentration at which the promoters become active (< 1 mM) is sufficient to maintain essential cellular processes permitting high-level protein synthesis (Table 4). Furthermore, their activity can be readily manipulated by limiting the P_i concentration (and glucose concentration for the *ugp* promoter) during the fermentation process. Their use should be applicable to the production of any of the categories of recombinant protein listed in Table 1.

Carbon-source-regulated promoters

The *mgl* promoter/operator region of the *mglBAEC* operon, which encodes a galactose transport system (Benner-Luger and Boos 1988), has been used to develop a series of expression vectors (Schumacher et al. 1988). Promoter activity is repressed by a protein encoded by the divergently transcribed *mglD* gene (Benner-Luger and Boos 1988) (Fig. 2). Repression is relieved by addition of D-fucose. The promoter is also

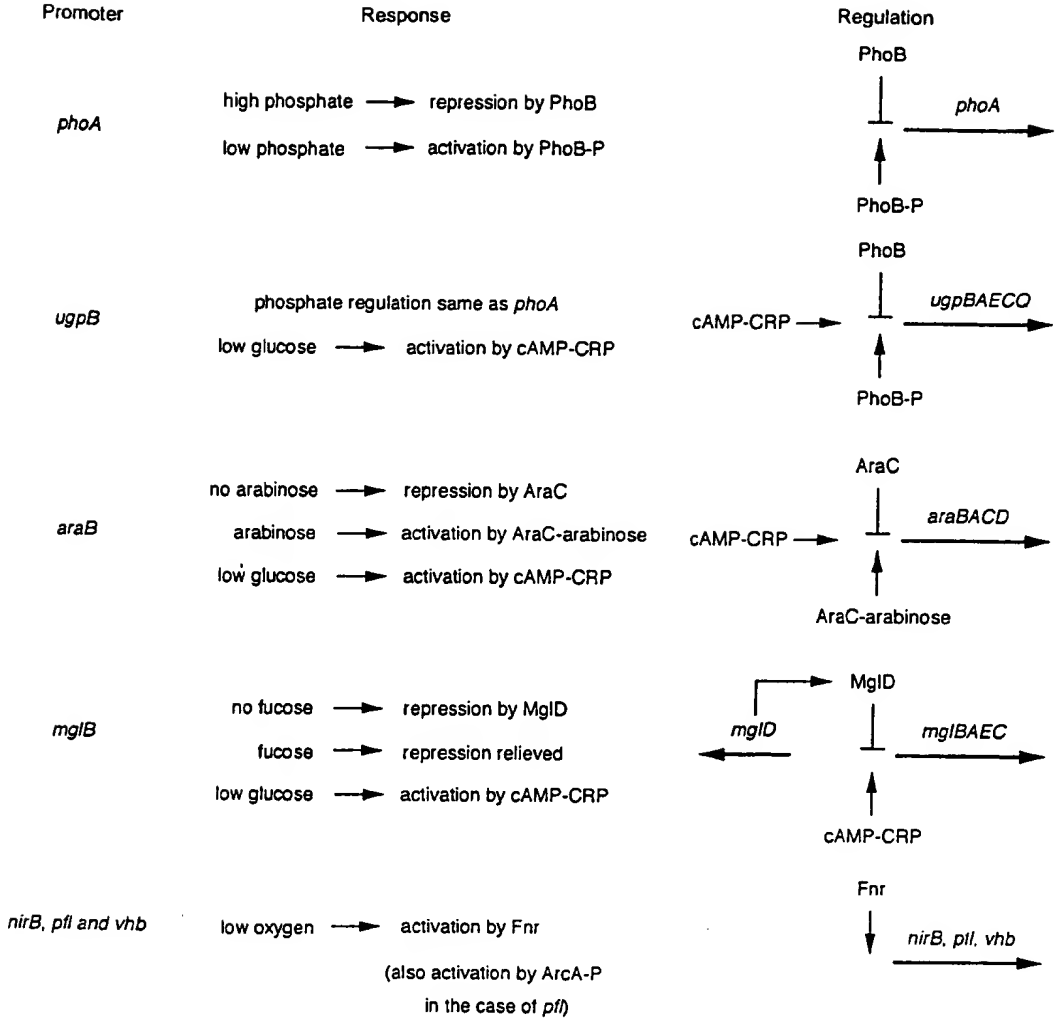


Fig. 2 Summary of the regulation of metabolically controlled promoters. → Genes or operons; ↑ activation; ⊥ repression of promoter activity

dependent on the cAMP-CRP complex for maximal activation and is therefore subject to glucose repression. Construction of a host strain and vector with the *mglD* gene deleted obviates the requirement for fucose addition and delivers an expression system that is controlled solely by catabolite repression. Induction of promoter activity is approximately 100-fold and this promoter has been used in an expression system that produced recombinant antigen from *Echinococcus multilocularis* at levels of 5 gl^{-1} (Müller et al. 1989) and more recently the glucose dehydrogenase from *Acinetobacter calcoaceticus*, also at concetrations of several grams per litre (M. Jarsch, unpublished results). Obviously the expression system regulated solely by glucose repression has the advantage that it can be readily and cheaply controlled in a large-scale process but it has the disadvantage that it can only be used for production of recombinant proteins that are both nontoxic to *E. coli* and have a background synthesis that does not reduce growth rates or cell yields extensively. Significantly, a recent study (Death and Ferenci 1994) has reported that when *E. coli* cells are grown in conditions of controlled glucose limitation sufficient endogenous inducer is synthesized to induce *mgl* operon expression to high levels. Optimization of the growth conditions should therefore permit the inclusion of the *mglD* gene in the vector, yielding an expression system with greater flexibility and concomitantly minimizing production costs by obviating the requirement for exogenous inducer.

A second sugar-regulated expression system utilizes the *araB* promoter from the arabinose operon of *Salmonella typhimurium* (Cagnon et al. 1991). The

activity of the *araB* promoter is subject to dual regulation through catabolite repression and induction by arabinose (Table 4). Promoter activity is controlled by the Ara C and CRP transcription factors (Fig. 2). Ara C represses transcription in the absence of arabinose but is a positive regulator of the promoter when it is itself complexed with D-arabinose (Lobell and Schleif 1991). D-Fucose is a structural analogue of arabinose and competes very efficiently with it to increase repression. Promoter activation, however, depends completely on the cAMP-CRP complex, even when arabinose is present (Lobell and Schleif 1991). It is possible to regulate the promoter over a 1200-fold range and vectors have been constructed that give yields of recombinant protein in the range of 30% of the total cellular protein (Cagnon et al. 1991). This system is similar to *lac*-promoter-based expression vectors in that glucose depletion and supply of an inducer form the basis of the regulation principle. The *araB* promoter has the added advantage that it can be switched off almost completely by fucose supplementation and hence provides a slightly more flexible system, in particular for production of toxic proteins. Moreover, arabinose is more cost-effective as an inducer compared with IPTG; on the basis of current market prices, IPTG is twice as expensive as D-arabinose if it is assumed that they are used at a final concentration of 5 mM and 1% (w/v) respectively.

Oxygen-regulated promoters

Promoters having an activity that can be modulated by varying the dissolved oxygen concentration of the culture medium offer several favourable advantages for the design of vectors to be used in the industrial production of recombinant proteins. First, expression is completely independent of the growth medium. Second, no addition of any inducing agents is necessary. Third, there is no dependence on a particular *E. coli* host strain. Fourth, the promoters are functional in the late exponential or early stationary phase of growth. Finally, high dissolved O_2 concentrations can be maintained until high cell densities are attained, and simply controlling the rate of aeration, the proportion of the oxygen in the gas mixture or regulating the speed of stirring can induce the expression of the recombinant gene. Since ensuring adequate aeration in high-density cultures is in any case a problem in fermenters, the use of oxygen-regulated promoters provides an inexpensive means of controlling product synthesis.

Three expression systems have been developed independently but all are based on the same regulation principle (Table 4, Fig. 2). The first system uses the promoter from the gene encoding bacterial haemoglobin (VHb). VHb is produced in large amounts by the gram-negative obligate aerobe *Vitreoscilla* when the dissolved O_2 becomes limiting (Khosla and Bailey 1989). The protein appears to function in oxygen delivery to the vigorously respiring membranes of the organism and it has been demonstrated that expression of the *vhb* gene in *E. coli* increased both the rate of growth and the cell densities attained (Khosla and Bailey 1988). The activity of the *vhb* promoter is regulated by oxygen in *E. coli* and it is likely that this may be controlled by the oxygen-responsive transcription factor Fnr (Spiro and Guest 1990). In a two-stage batch fermentation Khosla et al. (1990) overproduced chloramphenicol acetyltransferase and β-galactosidase, in independent constructs, to levels approaching 10% of the soluble cellular protein by growing recombinant *E. coli* to high cell densities at a dissolved O_2 concentration of 20% and then reducing the O_2 concentration to below 5% air saturation, which induced expression. An overall 30-fold increase in promoter activity was achieved (Khosla et al. 1990). Because the activity of the *vhb* promoter is optimal under conditions of microaerobiosis this means that the dissolved O_2 concentration must be carefully controlled to elicit high-level expression; the promoter has reduced activity when the culture becomes anaerobic (Khosla and Bailey 1989). It is noteworthy, however, that this promoter is functional over a broad host range, which includes *Pseudomonas*, *Azotobacter* and *Rhizobium* species (Dikshit et al. 1990).

We have developed an expression vector using the pyruvate formate-lyase operon regulatory region (Sawers and Böck 1989; Böck et al. 1990). Pyruvate formate-lyase is the central enzyme of anaerobic catabolism in *E. coli* and can reach levels exceeding 3% of the total cellular protein. The *pfl* gene is expressed anaerobically from multiple promoters and the induction ratio is maximally 30-fold (Table 4) (Sawers and Böck 1989). Anaerobic regulation is mediated by the Fnr and ArcA transcription factors (Sawers 1993). The vector that has been constructed incorporates two of the promoters plus the complete regulatory region and has been used successfully to produce in *E. coli* high levels of cholesterol oxidase from *Brevibacterium sterolicum* (Table 4). A disadvantage of this expression system is that the promoter has a significant basal activity when the organism is grown at high dissolved O_2 concentrations, therefore it is unsuitable if low background promoter activity is desired or necessary. Although the induction ratios of both the *pfl* and the *vhb* systems are moderate compared with the others listed, nevertheless they are based on strong promoters that allow high-level heterologous gene expression and their ease of control makes them very cost-effective.

Finally, a portion of the Fnr-dependent *nirB* promoter from the anaerobically inducible nitrite reeductase operon of *E. coli* has been used to produce an effective expression vector (Oxer et al. 1991). Like the *pfl* promoter *nirB* promoter activity is dependent on the Fnr transcription factor but it is also induced by nitrite. Oxer and colleagues (Oxer et al. 1991) removed the nitrite-responsive regulatory sequences but left the

recognition sequences for the Fnr protein intact. The resulting derivative exhibited tight transcriptional control in fermenter studies (Table 4) (Charles et al. 1992; Chatfield 1992). Two examples of heterologously produced proteins under the control of the *nirB* promoter have been described: tetanus toxin fragment C was produced at levels of 20% total cellular protein while the *Bordetella pertussis* antigen pertactin was synthesized at levels exceeding 30% of the total cellular protein (Oxer et al. 1991). A very low level of expression was observed when the oxygen supply was maintained, whereas anaerobiosis caused an approximately 100-fold induction of activity.

Conclusions and perspectives

The support of fundamental research has started to pay dividends in providing industry with new, strong, regulated expression systems that are suited to the production of recombinant proteins on a large scale in fermentation processes. Although they may not replace the established expression vectors for all purposes they offer attractive alternatives that have comparable promoter strength and, in the majority of cases, are easier and cheaper to control. The alternative promoters described, despite being strong, are not ferocious and experience using these systems to date has shown that not only is the yield of recombinant product high but also the fraction of insoluble, functionally inactive protein is generally quite low.

It is likely that future research will reveal further metabolically regulated promoters that may be suitable for the development of expression vectors. Indeed, a patent has been registered (Schroeckh et al. 1992) in which an expression vector has been developed that includes the *glnAp2* promoter from the glutamine synthetase operon of *E. coli*. The activity of the promoter is controlled by RNA polymerase containing an alternative sigma factor and is positively regulated by nitrogen deprivation. Alternative sigma factors confer upon RNA polymerase holoenzyme the ability to transcribe specifically promoters that have a different recognition sequence compared with the general sigma factor. A further class of regulated promoters, recognized by an alternative sigma factor, that could potentially be exploited for designing expression vectors includes those the activity of which is switched on only in the stationary phase of growth (Hengge-Aronis 1993). Finally, a series of vectors has recently been developed based on the *proU* promoter that is regulated in response to osmolarity (Herbst et al. 1994). The activity of the promoter is controlled by varying the concentration of NaCl in the growth medium and it exhibits excellent regulatory properties. These vectors therefore should prove extremely useful for the cost-effective production of recombinant proteins.

These promoters provide the basis for development of further novel expression vectors and, as our knowledge of the regulatory mechanisms underlying their control advances, it will perhaps be possible to combine features from different systems to develop hybrid or synthetic promoters with improved regulatory characteristics. The feasibility of this is shown by the development of hybrid promoters combining properties of established systems to deliver very efficient, tightly regulated expression vectors (Wülfing and Plückthun 1993; Chen et al. 1993).

Acknowledgements The work carried out by G.S. was performed in the Department of Microbiology at the University of Munich and was supported by the Deutsche Forschungsgemeinschaft.

References

Balbas P, Bolivar F (1990) Design and construction of expression plasmid vectors in *Escherichia coli*. Methods Enzymol 185:14–37

Beckwith J (1987) The lactose operon. In: Neidhardt FC, Ingraham JL, Low KB, Magasanik B, Schaechter M, Umbarger HE (eds) *Escherichia coli* and *Salmonella typhimurium*: molecular and cellular biology. American Society for Microbiology, pp 1444–1452

Benner-Luge D, Boos W (1988) The *mglB* sequence of *Salmonella typhimurium* LT2; promoter analysis by gene fusions and evidence for a divergently oriented gene coding for the *mgl* repressor. Mol Gen Genet 214:579–587

Böck A, Sawers G, Jarsch M, Herbst R (1990) Rekombinante DNA und Expressionsvektor. European patent application 0393684

Brosius J, Erfle M, Storella J (1985) Spacing of the −10 and −35 regions in the *tac* promoter. J Biol Chem 260:3539–3541

Cagnon C, Valverde V, Masson JL (1991) A new family of sugar-inducible expression vectors for *E. coli*. Protein Eng 4:843–847

Carter P, Kelley RF, Rodrigues ML, Snedecor B, Covarrubias M, Velligan MD, Wong WLT, Rowland AM, Kotts CE, Carver ME, Yang M, Bourell JH, Shepard, HM, Henner D (1992) High level *Escherichia coli* expression and production of a bivalent humanized antibody fragment. Biotechnology 10:163–167

Charles IG, Makoff AJ, Oxer MD, Peakman TC, Cole JA (1992) Production of recombinant proteins. Patent WO 9215688

Chatfield SN, Charles IG, Makoff AJ, Oxer MD, Dougan G, Pickard D, Slater D, Fairweather NF (1992) Use of the *nirB* promoter to direct the stable expression of heterologous antigens in *Salmonella* oral vaccine strains – development of a single-dose oral tetanus vaccine. Biotechnology 12:888–892

Chen W, Kallio PT, Bailey JE (1993) Construction and characterization of a novel crossregulation system for regulating cloned gene expression in *Escherichia coli*. Gene 130:15–22

Death A, Ferenci T (1994) Between feast and famine: endogenous inducer synthesis in the adaptation of *Escherichia coli* to growth on limiting carbohydrates. J Bacteriol 176:5101–5107

DeBoer HA, Comstock LJ, Vasser M (1983) The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. Proc Natl Acad Sci USA 80:21–25

Dikshit KL, Dikshit RP, Webster DA (1990) Study of *Vitreoscilla* globin (*vhb*) gene expression and promoter activity in *E. coli* through transcriptional fusion. Nucleic Acids Res 18:4149–4155

Fuller F (1982) A family of cloning vectors containing the *lacUV5* promoter. Gene 19:43–54

Gold L, Stormo GD (1990) High-level translation initiation. Methods Enzymol 185:89–93

Gottesman S (1990) Minimizing proteolysis in *Escherichia coli*: genetic solutions. Methods Enzymol 185:119–129

Gralla JD (1990) Promoter recognition and mRNA initiation by *Escherichia coli* $E\sigma^{70}$. Methods Enzymol 185:37–54

Gross G (1989) High-level gene expression in *Escherichia coli*. Chimicaoggi March:21–29

Hengge-Aronis R (1993) Survival of hunger and stress: the role of *rpoS* early stationary phase gene regulation in *E. coli*. Cell 72:165–168

Herbst B, Kneip S, Bremer E (1994) pOSEX: vectors for osmotically controlled and finely tuned gene expression *Escherichia coli*. Gene 151:137–142

Hockney RC (1994) Recent developments in heterologous protein production in *Escherichia coli*. Trends Biotechnol 12:456–463

Kasahara M, Makino K, Amemura M, Nakata A, Shinagawa H (1991) Dual regulation of the *ugp* operon by phosphate and carbon starvation at two interspaced promoters. J Bacteriol 173:549–558

Khosla C, Bailey JE (1988) Heterologous expression of a bacterial hemoglobin improves the growth properties of recombinant *Escherichia coli*. Nature 331:633–635

Khosla C, Bailey JE (1989) Characterization of the oxygen-dependent promoter of the *Vitreoscilla* hemoglobin gene in *Escherichia coli*. J Bacteriol 171:5995–6004

Khosla C, Curtis JE, Bydalek P, Swartz JR, Bailey JE (1990) Expression of recombinant proteins in *Escherichia coli* using an oxygen-responsive promoter. Biotechnology 8:554–558

Kopetzki E, Lehnert K, Buckel P (1994) Enzymes in diagnostics: achievements and possibilities of recombinant DNA technology. Clinical Chem 40:688–704

Lobell RB, Schleif RF (1991) AraC-DNA looping: orientation and distance-dependent loop breaking by the cAMP receptor protein. J Mol Biol 218:45–54

Makoff AJ, Oxer MD (1991) High level expression in *E. coli* using mutant forms of the *lac* promoter. Nucleic Acids Res 19: 2417–2421

Missiakis D, Georgopoulos C, Raina S (1993) Identification and characterization of the *Escherichia coli* gene *dsbB*, whose product is involved in the formation of disulfide bonds in vivo. Proc Natl Acad Sci USA 90:7084–7088

Müller N, Vogel M, Gottstein NB, Scholle A, Seebeck T (1989) Plasmid vector for overproduction and export of recombinant protein in *Escherichia coli*: efficient one-step purification of a recombinant antigen from *Echinococcus multilocularis* (Cestoda). Gene 75:329–334

Nygren P-A, Ståhl S, Uhlen M (1994) Engineering proteins to facilitate bioprocessing. Trends Biotechnol 12:184–188

Oka T, Sakamoto S, Miyoshi K, Fuwa T, Yoda K, Yamasaki M, Tamura G, Miyake T (1985) Synthesis and secretion of human epidermal growth factor by *Escherichia coli*. Proc Natl Acad Sci USA 82:7212–7216

Oxer MD, Bentley CM, Doyle JG, Peakman TC, Charles IG, Makoff IJ (1991) High level heterologous expression in *E. coli* using the anaerobically-activated *nirB* promoter. Nucleic Acids Res 19:2889–2892

Remaut E, Stanssans P, Fiers W (1981) Plasmid vectors for high-efficiency expression controlled by the p_L promoter of coliphage lambda. Gene 15:81–93

Richter G (1983) Glucose oxidase. In: Godfrey T, Reichelt J (eds) Industrial enzymology. The application of enzymes in industry. MacMillan, New York, pp 428–436

Rosenberg M, Ho YS, Shatzman A (1983) The use of pKC30 and its derivatives for controlled expression of genes. Methods Enzymol 101:123–138

Santos MAS, Tuite MF (1993) New insights into mRNA decoding – implications for heterologous protein synthesis. Trends Biotechnol 11:500–505

Sawers G (1993) Specific transcriptional requirements for positive regulation of the anaerobically inducible *pfl* operon by ArcA and FNR. Mol Microbiol 10:737–747

Sawers G, Böck A (1989) Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. J Bacteriol 171:2485–2498

Schroeckh V, MacCarthy JEG, Deckwer WD, Riesenberg D (1992) Expressionsvektor sowie Verfahren zur Expression rekombinater Gene unter Stickstoffmangelbedingungen. German patent DE 4102594

Schumacher G, Jarsch M, Boos W (1988) Expressionsvektor zur regulierbaren Expression von Fremdgenen in Prokaryonten. German patent application DE 3716957

Spiro S, Guest JR (1990) FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. FEMS Microbiol Rev 75:399–428

Squires CL, Lee FD, Yanofsky C (1975) Interaction of *trp* repressor and RNA polymerase with the *trp* operon. J Mol Biol 92:93–111

Stark MJR (1987) Multicopy expression vectors carrying the *lac* repressor gene for regulated high-level expression of genes in *Escherichia coli*. Gene 51:255–267

Studier FW, Moffat BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. J Mol Biol 189:113–130

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol 185:60–89

Su T-Z, Schweizer H, Oxender DL (1990) A novel phosphate-regulated expression vector in *Escherichia coli*. Gene 90:129–133

Su T-Z, Schweizer H, Oxender DL (1991) Carbon-starvation induction of the *ugp* operon, encoding the binding protein-dependent *sn*-glycerol-3-phosphate transport system in *Escherichia coli*. Mol Gen Genet 230:28–32

Tabor S, Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc Natl Acad Sci USA 82:1074–1078

Tacon W, Carey N, Emtage S (1980) The construction and characterization of plasmid vectors suitable for the expression of all DNA phases under the control of the *E. coli* tryptophan promoter. Mol Gen Genet 177:427–438

Wanner BL (1993) Gene regulation by phosphate in enteric bacteria. J Cell Biochem 51:47–54

Wanner BL, Wilmes MR, Young DC (1988) Control of bacterial alkaline phosphatase synthesis and variation in an *Escherichia coli* K-12 *phoR* mutant by adenyl cyclase, the cyclic AMP protein, and the *phoM* operon. J Bacteriol 170:1092–1102

Wülfing C, Plückthun A (1993) A versatile and highly repressible *Escherichia coli* expression system based on invertible promoters: expression of a gene encoding a toxic product. Gene 136: 199–203

Yansura DG, Henner DJ (1990) Use of the *Escherichia coli trp* promoter for direct expression of proteins. Methods Enzymol 185:54–60

THIS PAGE BLANK (USPTO)

Transgenic Research (1996) **5:** 49–55

Expression of a cystine-rich fish antifreeze in transgenic *Drosophila melanogaster*

BERNARD P. DUNCKER[1], J. ANNE HERMANS[2‡], PETER L. DAVIES[1,2] and VIRGINIA K. WALKER[1*]

Departments of Biology[1] and Biochemistry[2], Queen's University, Kingston, Ontario, K7L 3N6, Canada (Fax 1 613 545 6617)

Received 5 December 1994; revised 15 March 1995; accepted 16 March 1995

We have used *Drosophila melanogaster* as a model system for the transgenic expression of cystine-rich Type II antifreeze protein (AFP) from sea raven. This protein was synthesized and secreted into fly haemolymph where it migrated as a larger species (16 kDa) than the mature form of the protein (14 kDa) as judged by immunoblotting. *Drosophila*-produced Type II AFP demonstrated antifreeze activity both in terms of thermal hysteresis (0.13 °C) and inhibition of ice recrystallization. Recombinant AFP was purified and N-terminal sequencing revealed a 17 aa extension that began at the predicted signal peptide cleavage point. The expression of all three AFP types in transgenic *Drosophila* has now been achieved. We conclude that the globular Type II and Type III AFPs are better choices for antifreeze transfer to other organisms than is the more widely used linear Type I AFP.

Keywords: proprotein; secretion; haemolymph; thermal hysteresis

Introduction

Organisms that inhabit freezing environments have developed a number of approaches to protect their body fluids from unrestrained ice formation (Storey and Storey, 1988). One such strategy is to synthesize antifreeze proteins (AFPs). Although structurally diverse, AFPs from different species act in a similar fashion (Davies and Hew, 1990). They all inhibit freezing by binding to nascent ice crystals, thereby making the addition of water to the ice lattice less favourable (DeVries, 1983). This results in a non-colligative lowering of the fluid freezing point below the melting point, which is known as thermal hysteresis. AFPs also inhibit ice recrystallization (Knight, *et al.*, 1984). This is important for freeze-tolerant organisms because it reduces the size of ice crystals that form upon thawing, and thus limits the damage done to surrounding tissues (Knight and Duman, 1986).

The transgenic expression of AFPs offers promise as a means of conferring freeze resistance to species that do not normally synthesize proteins with thermal hysteresis activity. To date, antifreeze genes have only been isolated from fish, but some of these genes, or their synthetic variants, have already been used in transgenic applications. AFP-producing transgenic tobacco lines have been generated in the hopes of conferring frost resistance (Hightower *et al.*, 1991; Kenward *et al.*, 1993), and AFP genes have also been transferred to salmon with the aim of increasing their survival in ice-laden seawater (Shears *et al.*, 1991). While the AFP expression in these commercially important species is encouraging, measurable resistance to freezing has yet to be achieved. We have adopted the fruit fly *Drosophila melanogaster* as a model system to evaluate the transgenic expression of fish AFP genes. The alanine-rich Type I AFP of winter flounder was the first type to be produced in flies (Rancourt *et al.*, 1987). Although synthesis of this protein was detected on immunoblots, AFP levels were insufficient to register thermal hysteresis activity in the haemolymph. Subsequently, this protein has been produced in transgenic tobacco where levels of accumulation have been equally disappointing (Kenward *et al.*, 1993). The failure of the protein to accumulate has been attributed in part to its lack of tertiary structure and a resultant susceptibility to proteolysis.

*To whom correspondence should be addressed.

‡Present address: Agriculture Canada, Siebens-Drake Research Institute, 1400 Western Road, London, ON, N6G 2V4, Canada

0962–8819 © 1996 Chapman & Hall

Far greater success was achieved when the Type III AFP gene from Atlantic wolffish was expressed in *Drosophila* (Rancourt *et al.*, 1990). This gene codes for a tightly structured, globular protein that has no particular amino acid bias (Chao *et al.*, 1993; Sönnichsen *et al.*, 1993). Levels of circulating AFP in excess of 1 mg ml^{-1} were obtained and resulted in haemolymph thermal hysteresis values of up to 0.35 °C.

We have investigated a third candidate for transgenic expression, the Type II AFP gene of sea raven. Its antifreeze is characteristically rich in cysteinyl residues, all of which form disulfide bridges in the folded protein (Ng and Hew, 1992). In fish, Type II AFP is initially synthesized as a 163 aa protein, but the circulating form has 129 aa residues (Hayes *et al.*, 1989). Current signal peptide cleavage algorithms (von Heijne, 1986) predict an intermediate of 146 aa, and a protein of this length is produced in fall armyworm cells transformed with recombinant baculovirus containing Type II AFP cDNA (Duncker *et al.*, 1994). In this report we examine the synthesis, processing, and activity of Type II AFP in the whole animal model system and compare its suitability for transgenic applications to that of the other two AFP types.

Materials and methods

AFP expression construct P[w, YP1:srAFP]

A 550 bp *Hin*dIII-*Pst*I fragment (Fig. 1b) derived from sea raven AFP cDNA clone C2-1 (Ng *et al.*, 1986) was ligated into pTZ19R (Pharmacia). To facilitate a subsequent *Bgl*I/*Bgl*I ligation step, the *Bgl*I sequence at position +133 relative to the cap site was altered by site-directed mutagenesis from GCCATGATGGC to an alternate *Bgl*I sequence, GCCACCTTGGC. A fusion was then made between the mutagenized sea raven AFP cDNA coding for exons 1–4 and the genomic AFP DNA coding for intron 4 and exon 5 (Fig. 1c). This was achieved by the three-way directional ligation of (i) the 550 bp *Hin*dIII/*Pst*I cDNA fragment; (ii) the 640 bp *Eco*RI/*Pst*I fragment (Fig. 1a) of the 2.4 kb *Eco*RI/*Hin*dIII subclone of genomic clone SR7 (Hayes *et al.*, 1989); into (iii) *Eco*RI/*Hin*dIII-cut pUC19 to give psrAFP C/G. The 1639 bp *Bam*HI/*Pst*I fragment from pDMYP1/2:IR^a (Riddell *et al.*, 1981) was ligated into pUC9, generating pYP1,2 (BP) (not shown). The 370 bp *Bgl*I/*Nsi*I YP fragment from pYP1,2 (BP) and the 1 kb *Bgl*I/*Eco*RI AFP fragment from psrAFP C/G were directionally ligated into the *Eco*RI and *Pst*I sites of the pW8 *P*-element vector polylinker (Klemenz *et al.*, 1987), to generate P[w, YP1:srAFP] (Fig. 1d). Although they have compatible cohesive ends for ligation, neither the *Pst*I site nor the *Nsi*I site is regenerated. This final ligation was facilitated through the earlier mutagenesis of the AFP *Bgl*I site to make the two *Bgl*I termini compatible.

Embryo microinjections and fly husbandry

Pre-syncytial blastoderm embryos collected from the flightless, white-eyed *Drosophila* host strain yw^{67};Ifm(3)3 (Duncker *et al.*, 1993) were microinjected with a mixture of P[w, YP1:srAFP] and transposase expression vector pπ25.7 wc (Spradling, 1986) in injection buffer (5 mM KCl, 0.1 mM NaH_2PO_4, pH 6.8), at concentrations of 350 ng λl^{-1} and 50 ng λl^{-1}, respectively, following standard procedures (Rubin and Spradling, 1982). Surviving G_0 flies were mated with injection stock flies, and the red eye colour of resulting progeny was used to identify transgenic G_1 flies. Homozygous lines were obtained by monitoring the eye phenotype through single pair matings.

Collection of fly haemolymph

Newly emerged (<24 h) adult flies were collected and placed on fresh food (10% yeast–sucrose medium) (Walker, 1985) for 48 h to ensure maximal *yp1* promoter-directed expression (Bownes *et al.*, 1988). For small-scale isolations of haemolymph, 50–100 flies were placed in 1.5 ml microcentrifuge tubes and frozen in dry ice for 5–10 min. For large-scale isolations, 1000–5000 flies were frozen in 50 ml centrifuge tubes. Following vigorous shaking to detach heads and legs, the flies were transferred to either 1 ml pipette tips (small-scale) or 5 ml syringes (large-scale) plugged with glass wool and were centrifuged (10 min, 10 000 × g, 4 °C) to collect the extruded haemolymph. PMSF was added to all samples at a final concentration of 5 mM. Haemocytes were pelleted by a second round of centrifugation (5 min, 14 000 × g, 4 °C) and the supernate was transferred to a fresh tube which was stored at −20 °C until needed.

Antifreeze activities

Ice crystal morphology and thermal hysteresis measurements were conducted using a nanolitre osmometer (Clifton Technical Physics, Hartford, New York, USA) following standard procedures (Chakrabartty and Hew, 1991). Photographs were taken with a WILD MPS 12 microcamera attachment through a Leitz Dialux 22 stereomicroscope, (Wild Heerbrugg, Heerbrugg, Switzerland). Thermal hysteresis is defined as the difference between melting and non-equilibrium freezing temperatures in °C. Ice recrystallization was observed by cryomicroscopy (Carpenter and Hansen, 1992).

Protein purification

Haemolymph (9 ml) collected from flies (50 g) of transgenic line P[w, YP1:srAFP]1 was applied to a Sephadex G-75 (Pharmacia) column (100 cm × 2.6 cm) equilibrated with 5 mM Tris-HCl, (pH 9.0). Aliquots of fractions were lyophilized, resuspended in 0.1 M

NH_4HCO_3 and assayed for thermal hysteresis activity. Pooled, active fractions were chromatographed on a Mono-Q FPLC column (Pharmacia) using a NaCl gradient in 5 mM Tris-HCl, (pH 9.0), at a flow rate of 0.5 ml min^{-1}. The NaCl gradient increased linearly in increments from 0 to 0.1 M in 10 min, from 0.1 M to 0.25 M in 30 min, and from 0.25 M to 0.5 M in 10 min. Fractions (0.5 ml) corresponding to the major Mono-Q peak were pooled and rechromatographed by reversed-phase HPLC on a C18 column (Vidac), using an acetonitrile gradient in 0.1% trifluoroacetic acid, at a flow rate of 1 ml min^{-1}.

Protein sequencing

Automated Edman degradations were done using an Applied Biosystems pulsed liquid sequencer (model 473A) equipped with microgradient phenylthiohydantoin analysis.

Results

Type II AFP is produced as a proprotein in transgenic Drosophila

The chimeric gene, P[w, YP1:srAFP], was constructed to place a sea raven AFP cDNA/genomic hybrid under the transcriptional control of the female-specific *Drosophila ypl* promoter (Fig. 1). Genomic sequence was incorporated to include an intron in the transcription unit because we have found that the presence of an intervening sequence boosts levels of Type III AFP transcripts in transgenic *Drosophila* (unpublished results). Two transgenic fly lines, P[w, YP1:srAFP]1 and P[w, YP1:srAFP]2, were generated through embryo microinjections with P[w, YP1:srAFP]. Homozygous stocks were established and the integrity of their transgenes was confirmed by Southern analysis (Fig. 2). Three hybridization signals were observed on DNA blots of both transgenic lines. A common fragment at 2.3 kb, which was also present in the injection stock, represented endogenous *ypl*. A second common fragment, at 0.9 kb, confirmed the integrity of the internal *Pvu* II fragment of the insert. Both transgenic lines displayed an additional unique fragment, the size of which was determined by the location of the nearest genomic *Pvu* II site downstream of the different integration sites of the transferred vector.

Haemolymph collected from P[w, YP1:srAFP]1 and P[w, YP1:srAFP]2 flies was subjected to SDS-PAGE and transblotted onto a nylon membrance (Millipore Immobilon-P). Immunoblots using rabbit anti-sea raven AFP antiserum (Ng *et al.*, 1986) revealed a cross-reacting 16 kDa protein in the haemolymph from female flies of both transgenic lines (Fig. 3, lanes b and d). No such

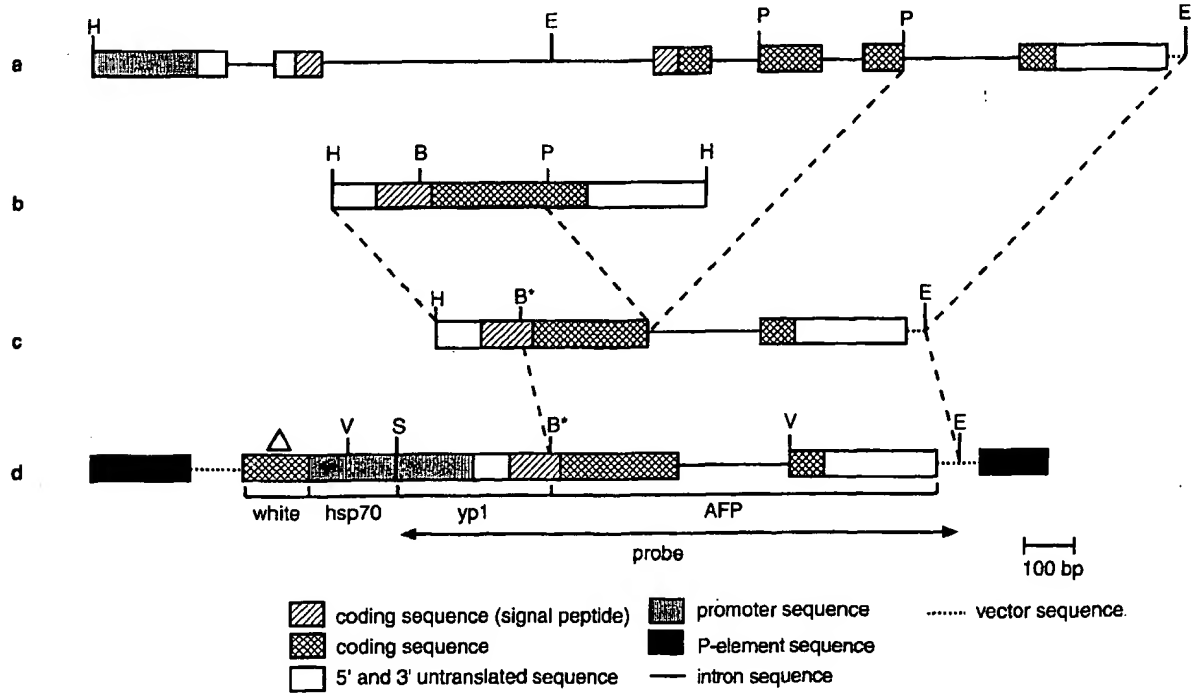


Fig. 1. Construction of sea raven AFP expression vectors. (a) sea raven AFP gene; (b) sea raven AFP cDNA; (c) cDNA/genomic fusion clone psrAFP C/G; (d) P-element expression construct P[w, YP1:srAFP]. Genes and constructs are drawn to scale, and pertinent restriction sites *Bgl* I (B), *Eco* RI (E), *Hin* dIII (H), *Pst* I (P), *Pvu* II (V) and *Sph* I (S) are indicated. The mutagenized *Bgl* I site is denoted by an asterisk. The double-headed arrow indicates the fragment (*Sph* I/*Eco* RI) used as a probe for Southern analysis. The triangle signifies that the white gene is not drawn to scale. Further details of the construction are presented in Materials and methods.

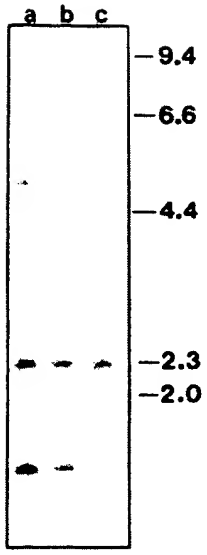


Fig. 2. Southern analysis of transgenic fly lines. Genomic DNA (10–20 λg) was isolated from flies of each line (Chia *et al.*, 1985) and digested with *Pvu* II. DNA fragments were separated by agarose gel electrophoresis and transferred to a nylon membrane. The fragment indicated in Fig. 1 was nick-translated and used as a probe. Lanes (a) and (b), DNA from transgenic lines, P[w, YP1:srAFP]1 and P[w, YP1:srAFP]2, respectively. Lane (c), DNA from the injection stock yw^{67};Ifm(3)3.

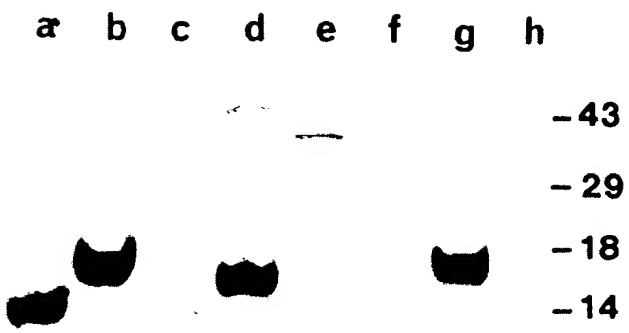


Fig. 3. Immunoblot analysis of insect-produced AFP. SDS-PAGE was performed using a 15% polyacrylamide SDS gel, containing 0.1 M sodium phosphate (pH 6.8) and 4 M urea. Prior to loading, all samples were boiled in loading buffer (0.1% SDS, 0.1 M sodium phosphate, 1% β-mercaptoethanol, 8 M urea, 0.003% bromophenol blue) for 5 min. Transfer to a nylon membrance was carried out. The membrane was incubated with rabbit anti-sea raven AFP antiserum and then with horseradish peroxidase-linked goat anti-rabbit IgG. Detection was performed using enhanced chemiluminescence (Amersham). Samples: purified AFP from sea raven serum (a); haemolymph (0.1 λl) from P[w, YP1:srAFP]1 females (b) and males (c); P[w, YP1:srAFP]2 females (d) and males (e); yw^{67};Ifm(3)3 injection stock females (f); culture medium (15 λl) from recombinant (g) and wild-type (h) baculovirus-infected fall armyworm cells. The position and size (kDa) of the protein standards are indicated at the side. The slight mobility difference between the observed AFPs in lanes b and d was an electrophoretic artifact of this particular gel. Additional high M_r bands are commonly, but not consistently seen for haemolymph samples challenged with this antiserum (compare lanes d and e to lanes b,c and f). It is not known what these bands represent.

protein was evident in the haemolymph of male transgenics (lanes c and e) or in that of nontransgenic, injection stock females (lane f). Mature Type II AFP purified from sea raven serum migrated as a 14 kDa protein (lane a). Culture medium from AFP-producing recombinant baculovirus-infected fall armyworm cells (lane g) as well as medium from wild-type baculovirus-infected cells (lane h) (Duncker *et al.*, 1994) were used as a standard and control, respectively. The baculovirus- and *Drosophila*-produced Type II AFPs demonstrated a similar mobility.

Recombinant proAFP demonstrates antifreeze activity

Upon cooling, ice crystals in the haemolymph samples from female P[w, YP1:srAFP]1 transgenics grew into hexagonal bipyramids (Fig. 4b). These were stable in size and, as noted by the slight curvature at the junction of the two hexagonal pyramids, characteristic of those seen in Type II AFP solutions from fish (Fig. 4a). In contrast, those observed in haemolymph samples from host nontransgenic female flies (yw^{67};Ifm(3)3) lacked this bipyramidal shape and demonstrated unrestrained growth (Fig. 4c). The thermal hysteresis value for female P[w, YP1:srAFP]1 haemolymph was determined to be 0.13 ± 0.2 °C.

These same haemolymph samples were evaluated for their ability to inhibit ice recrystallization. After snap

Fig. 4. Ice crystal morphology. Samples analysed were sea raven AFP purified from fish serum (a), haemolymph from female P[w, YP1:srAFP]1 (b) and female yw^{67};Ifm(3)3 (c) flies. Magnification was 700-fold.

freezing to −133 °C, the ice crystal size at a number of temperatures approaching the haemolymph melting point was recorded. When compared to haemolymph from nontransgenic flies, P[w, YP1:srAFP]1 haemolymph

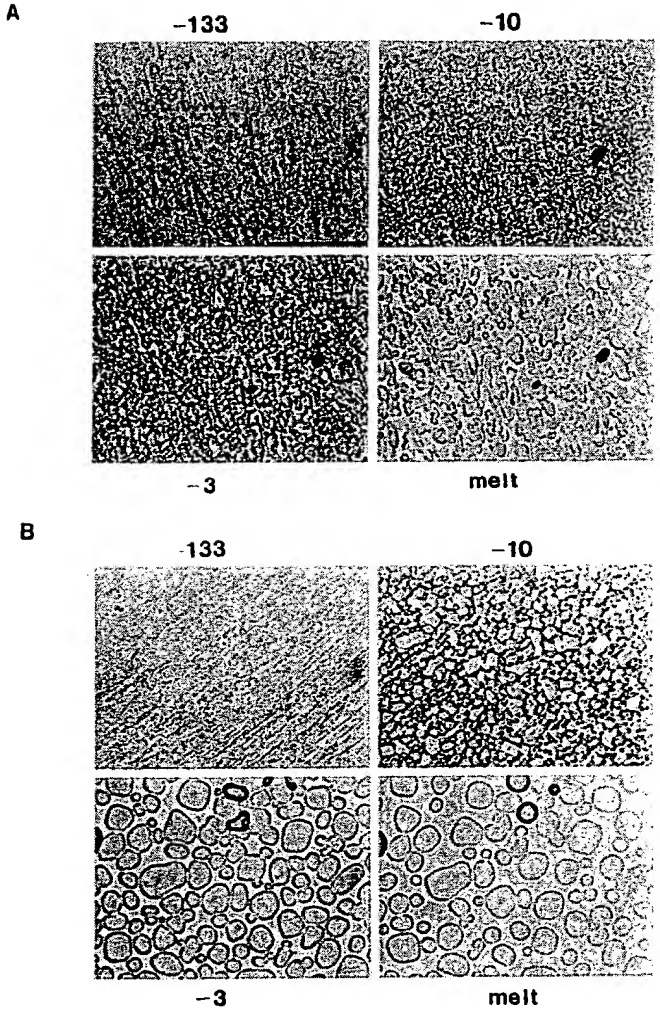


Fig. 5. Ice recrystallization assays. Haemolymph from female P[w, YP1:srAFP]1 (panel A) and yw^{67};lfm(3)3 (panel B) flies was snap frozen at −133 °C and micrographs were taken at −133 °C, −10 °C, −3 °C, and at the onset of melting. The bar represents 50 λm; all micrographs are to the same magnification.

formed smaller ice crystals at −10 °C, −3 °C and at the onset of melting (Fig. 5), indicating inhibition of ice recrystallization.

Purification and characterization of Drosophila-produced Type II proAFP

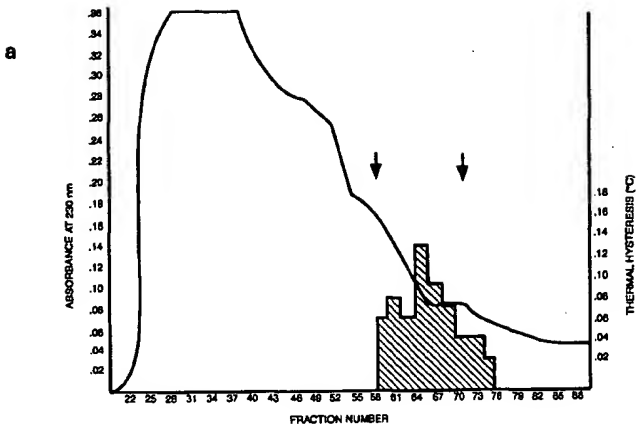
AFP from P[w, YP1:srAFP]1 haemolymph was purified by Sephadex G-75 gel permeation chromatography (Fig. 6a) followed by Mono-Q FPLC (Fig. 6b) and reversed-phase HPLC (Fig. 6c). The AFP was detected by thermal hysteresis activity in the low M_r end of the G-75 eluate, consistent with a M_r of 16 000. On the Mono-Q column the activity co-chromatographed with the major A_{280}-absorbing peak but based on the subsequent HPLC profile was only a very minor constituent of it. Purified recombinant sea raven AFP was subjected to 10 cycles of automated Edman degradation and gave a N-terminal sequence of NDDKILKGTA. This matches the expected N-terminus for the proform of the recombinant AFP beginning at residue 18 of the 163 aa primary translation product, as predicted by the von Heijne rules for signal peptide cleavage (von Heijne, 1986).

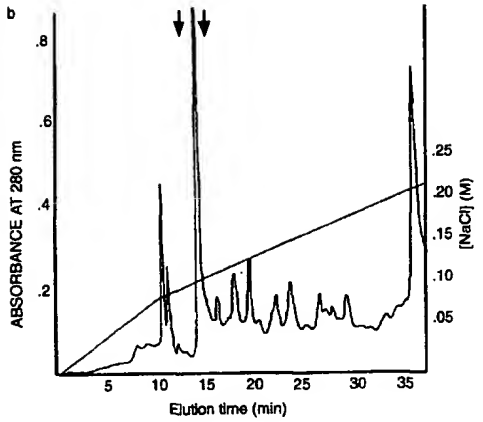
Discussion

We have produced active Type II AFP in transgenic *Drosophila melanogaster*. To ensure efficient secretion into the haemolymph, the fusion between *yp1* and AFP gene sequences was made in the regions encoding their signal peptides in such a way as to preserve a hydrophobic core within the signal sequence. The YP1 signal sequence was changed from MNPMRVLSLLACLAVAALAK to MNPMRVLSLLACLAVAALA↓**LTQA**. The fusion point is indicated by the arrow. Using the rules of von Heijne (1986), signal peptide cleavage is predicted to occur after the Ala, four residues C-terminal to the fusion point of the hybrid signal peptide (above), just as it does in the native fish signal peptide. This strategy for designing a chimaeric signal sequence had previously been used for the correct processing and secretion of Type III AFP in transgenic *Drosophila* (Rancourt *et al.*, 1990).

As predicted, a 146 aa proprotein was produced in the flies. Since there was no further processing of the proAFP, it is likely that the proteinase activity responsible for pro region removal in sea raven is absent in *Drosophila*. The antifreeze activity of the haemolymph from transgenic flies was demonstrated by both thermal hysteresis and inhibition of ice recrystallization. Haemolymph from control flies lacked these activities. Since Type II AFP loses its activity under reducing conditions (Slaughter *et al.*, 1981) it appears that it is correctly folded through disulfide bridge formation. These observations are consistent with our previous experience using insect cell culture where an active proprotein was also synthesized (Duncker *et al.*, 1994).

This now represents the third antifreeze type to be produced in transgenic flies. Previously, winter flounder (Type I) and wolffish (Type III) AFP genes were used in expression constructs where each was placed under the transcriptional control of *Drosophila* yolk protein promoters. Type I could not be detected in transgenic flies (Rancourt *et al.*, 1992), although this protein was transiently synthesized under heat shock conditions when its gene was linked to the *Drosophila hsp70* promoter (Rancourt *et al.*, 1987). In contrast, fly lines carrying the Type III gene synthesized AFP levels that gave freezing point depressions of 0.1–0.35 °C. The fly lines in this report, P[w, YP1:srAFP]1 and 2, produced haemolymph thermal hysteresis readings of up to 0.13 °C. Type II and





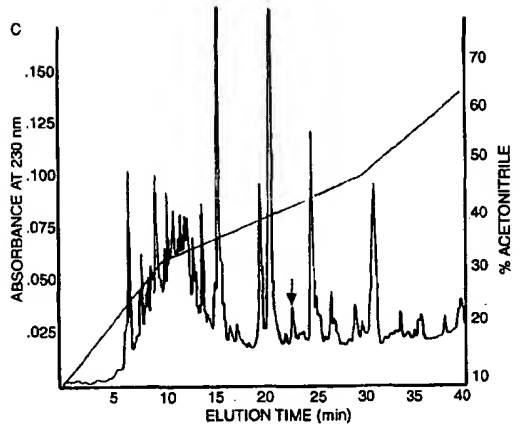


Fig. 6. Purification of *Drosophila*-produced AFP from line P[w, YP1:srAFP]1. (a) Sephadex G-75. Absorbance readings at 230 nm are indicated by the tracing while thermal hysteresis (freezing point depression) is represented by the histogram. Fractions between the arrows were pooled. (b) Ion exchange FPLC. Pooled G-75 fractions were chromatographed on a Mono-Q column (Pharmacia). Fractions between the arrows demonstrated antifreeze activity and were pooled (1.5 ml) for further purification. (c) Reversed-phase HPLC. Pooled FPLC fractions were loaded onto the HPLC column in solution A (0.1% trifluoroacetic acid) and eluted by a gradient of solution B (80% acetonitrile in 0.1% trifluoroacetic acid). The solution B concentration was increased linearly to 40% over 10 min, to 60% over a further 20 min, and to 80% over a subsequent 10 min. Fractions (1 ml) were collected. The greatest antifreeze activity was detected in the fraction corresponding to the peak indicated by the arrow (0.74 °C, for a tenfold concentration in 0.1 M NH_4HCO_3). This was the only peak showing significant thermal hysteresis. The large number of additional peaks likely represent oxidized haemolymph as well as melanin adducts and polymers from the insect wound response (Boman and Hultmark, 1987). Further details of purification steps are presented in Materials and methods.

III thermal hysteresis levels were therefore reasonably comparable, bearing in mind that the Type III expression construct included two copies of the AFP gene, whereas the Type II construct used in this study had only one. These results indicate that sea raven Type II AFP as well as the Type III AFP of wolffish are much better candidates in developing freeze resistance than the more widely used flounder Type I AFP (Cutler *et al.*, 1989; Georges *et al.*, 1990; Hightower *et al.*, 1991; Shears *et al.*, 1991; Kenward *et al.*, 1993). One possible reason for this difference is that Type I AFP has no tertiary structure and is less stable than the globular AFPs at room temperature; at −1 °C flounder AFP is 87% α-helical whereas at 25 °C the proportion drops to 47% (Ananthanarayanan and Hew, 1977). The lower α-helical content at room temperature might therefore contribute to a faster rate of degradation. In studies where a winter flounder Type I AFP gene was expressed in tobacco, AFP was detectable when the plants were grown at 4 °C, but not when they were placed at room temperature (Kenward *et al.*, 1993). In experiments with transgenic *Drosophila* expressing flounder AFP, we have found that this protein persists for much longer when flies are reared at 10 °C, rather than at room temperature (unpublished results, 1994).

The insights gained using this system have demonstrated its potential to assess candidate AFPs for use in other transgenic organisms. Additionally, it may have applications in the study of insect AFPs that, like the sea raven Type II AFP, appear to be extensively disulfide-bonded.

Acknowledgements

We would like to thank Dr Choy Hew (University of Toronto) for the gift of anti-sea raven AFP antisera, and

Dr John Carpenter (University of Colorado Health Science Centre) for cryomicroscopy. This research was supported by MRC and NSERC grants to P.L.D and V.K.W., respectively, and by an Ontario Graduate Scholarship awarded to B.P.D.

References

Ananthanarayanan, V.S. and Hew, C.L. (1977) Structural studies on the freezing point-depressing protein of the winter flounder *Pseudopleuronectes americanus*. *Biochem. Biophys. Res. Commun.* **74**, 685–9.

Boman, H.G. and Hultmark, D. (1987) Cell-free immunity in insects. *Ann. Rev. Microbiol.* **41**, 103–26.

Bownes, M., Scott, A. and Shirras, A. (1988) Dietary components modulate yolk protein gene transcription in *Drosophila melanogaster*. *Development* **103**, 119–28.

Carpenter, J.F. and Hansen, T.N. (1992) Antifreeze protein modulates cell survival during cryopreservation: mediation through influence on ice crystal growth. *Proc. Natl Acad. Sci. USA* **9**, 8953–7.

Chakrabartty, A. and Hew, C.L. (1991) The effect of enhanced alpha-helicity on the activity of a winter flounder polypeptide. *Eur. J. Biochem.* **202**, 1057–63.

Chao, H., Davies, P.L. Sykes, B.D. and Sönnichsen, F.D. (1993) Use of proline mutants to help solve the NMR solution structure of type III antifreeze protein. *Protein Sci.* **2**, 1411–28.

Chia, W., Karp, R., McGill, S. and Ashburner, M. (1985) Molecular analysis of the *Adh* region of the genome of *Drosophila melanogaster*. *J. Mol. Biol.* **186**, 689–706.

Cutler, A.J., Saleem, M., Kendall, E., Gusta, L.V., Georges, F. and Fletcher, G.L. (1989) Winter flounder antifreeze protein improves the cold hardiness of plant tissues. *J. Plant Physiol* **135**, 351–4.

Davies, P.L. and Hew, C.L. (1990) Biochemistry of fish antifreeze proteins. *FASEB J.* **4**, 2460–8.

DeVries, A.L. (1983) Antifreeze peptides and glycopeptides in cold-water fishes. *Ann. Rev. Physiol* **45**, 245–60.

Duncker, B.P., Hermans, J.A., Davies, P.L. and Walker, V.K. (1993) Biological containment of transgenic fruit files. *Dros. Info. Serv.* **72**, 133.

Duncker, B.P., Gauthier, S.Y. and Davies, P.L. (1994) Cystine-rich fish antifreeze is produced as an active proprotein precursor in fall armyworm cells. *Biochem. Biophys. Res. Commun.* **203**, 1851–7.

Georges, G., Saleem, M. and Culter, A.J. (1990) Design and cloning of a synthetic gene for the flounder antifreeze protein and its expression in plant cells. *Gene* **91**, 159–65.

Hayes, P.H., Scott, G.K., Ng, N.F., Hew, C.L. and Davies, P.L. (1989) Cystine-rich type II antifreeze protein precursor is initiated from the third AUG codon of its mRNA. *J. Biol. Chem.* **264**, 18761–7.

Heijne, G. von (1986) A new method for predicting signal peptide sequence cleavage sites. *Nucl. Acids Res.* **14**, 4683–91.

Hightower, R., Baden, C., Penzes, E., Lund, P. and Dunsmuir, P. (1991) Expression of antifreeze proteins in transgenic plants. *Plant Mol. Biol.* **17**, 1013–21.

Kenward, K.D., Altschuler, M., Hildebrand, D. and Davies, P.L. (1993) Accumulation of type I fish antifreeze protein in transgenic tobacco is cold-specific. *Plant Mol. Biol.* **23** 377–85.

Klemenz, R., Weber, U. and Gehring, W.J. (1987) The white gene as a marker in a new P-element vector for gene transfer in *Drosophila*. *Nucl. Acids Res.* **15**, 3947–59.

Knight, C.A. and Duman, J.G. (1986) Inhibition of recrystallization of ice by insect thermal hysteresis proteins: a possible cryoprotective role. *Cryobiology* **23**, 256–62.

Knight, C.A., Hallett, J. and De Vries, A.L. (1984) Fish antifreeze protein and the freezing and recrystallization of ice. *Nature* **308**, 295–6.

Ng, N.F.L. and Hew, C.L. (1992) Structure of an antifreeze polypeptide from the sea raven. *J. Biol. Chem.* **267**, 16069–75.

Ng, N.F.L., Khiet-Yan, T. and Hew, C.L. (1986) Structure of an antifreeze polypeptide precursor from the sea raven, *Hemitripterus americanus*. *J. Biol. Chem.* **261**, 15690–5.

Rancourt, D.E., Walker, V.K. and Davies, P.L. (1987) Flounder antifreeze protein synthesis under heat shock control in transgenic *Drosophila melanogaster*. *Mol. Cell. Biol.* **7**, 2188–95

Rancourt, D.E., Peters, I.D., Walker, V.K. and Davies, P.L. (1990) Wolffish antifreeze protein from transgenic *Drosophila*. *Bio/Technology* **8**, 453–7.

Rancourt, D.E., Davies, P.L. and Walker, V.K. (1992) Differential translatability of antifreeze protein mRNAs in a transgenic host. *Biochim. Biophys. Acta* **1129**, 188–94.

Riddell, D.C., Higgins, M.J., McMillan, B.J. and White, B.N. (1981) Structural analysis of three vitellogenin genes in *Drosophila melanogaster*. *Nucl. Acids Res.* **9**, 1323–38.

Rubin, G.M. and Spradling, A.C. (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–53.

Shears, M.A., Fletcher, G.L., Hew, C.L., Gauthier, S. and Davies, P.L. (1991) Transfer, expression and stable inheritance of antifreeze protein genes in atlantic salmon (*Salmo salar*). *Mol. Mar. Biol. Biotech* **1**, 58–63.

Slaughter, D., Fletcher, G.L., Ananthanarayanan, V.S. and Hew, C.L. (1981) Antifreeze proteins from the sea raven, *Hemitripterus americanus*. Further evidence for diversity among fish polypeptide antifreezes. *J. Biol. Chem.* **256**, 2022–6.

Sönnichsen, F.D., Sykes, B.D., Chao, H. and Davies, P.L. (1993) The nonhelical structure of antifreeze protein Type III. *Science* **259**, 1154–7.

Spradling, A.C. (1986) P-element mediated transformation. In: Roberts, D.B. (ed.) *Drosophila: A Practical Approach*. pp. 175–97 Oxford: IRL Press.

Storey, K.B. and Storey, J.M. (1988) Freeze tolerance in animals. *Physiol. Rev.* **68**, 27–84.

Walker, V.K. (1985) Yolk polypeptide gene expression in Minute *Drosophila* females. *Biochem. Genet.* **23**, 363–78.

THIS PAGE BLANK (USPTO)

Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 7036–7040, July 1995
Genetics

Green fluorescent protein as a vital marker and reporter of gene expression in *Drosophila*

(targeted expression/enhancer trap/GAL4/confocal microscopy)

EDWARD YEH*, KERSTIN GUSTAFSON*, AND GABRIELLE L. BOULIANNE†

Centre for Research in Neurodegenerative Diseases, Department of Physiology and Zoology, Tanz Neuroscience Building, University of Toronto, Toronto, ON, Canada M5S 1A8

Communicated by Allan C. Spradling, Carnegie Institution of Washington, Baltimore, MD, April 21, 1995 (received for review March 16, 1995)

ABSTRACT **We have used the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* as a vital marker/reporter in *Drosophila melanogaster*. Transgenic flies were generated in which GFP was expressed under the transcriptional control of the yeast upstream activating sequence that is recognized by GAL4. These flies were crossed to several GAL4 enhancer trap lines, and expression of GFP was monitored in a variety of tissues during development using confocal microscopy. Here, we show that GFP could be detected in freshly dissected ovaries, imaginal discs, and the larval nervous system without prior fixation or the addition of substrates or antibodies. We also show that expression of GFP could be monitored in intact living embryos and larvae and in cultured egg chambers, allowing us to visualize dynamic changes in gene expression during real time.**

Development is the cumulative effect of dynamic changes in gene expression in different cells within an organism. At present, several techniques exist that allow an examination of gene expression through the measurement of either RNA or protein distribution within fixed tissue. Gene expression can be measured either directly by using probes and antibodies or indirectly by detecting the product of a fusion between the gene of interest and a reporter gene such as bacterial *lacZ* (1). In *Drosophila*, *lacZ* is often used in enhancer trap screens to identify genes that are expressed in a tissue-specific manner (2–4) or as a reporter to identify tissue-specific regulatory regions within known genes. All of these approaches are limited in that they only provide a static image of changes in gene expression during development. Furthermore, these techniques usually involve extensive manipulation including dissection of the tissue of interest, fixation, and the addition of various substrates or antibodies, and they are of limited use in living tissue.

Recently, Chalfie *et al.* (5) described the use of the green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* as a vital reporter for gene expression in both bacteria and *Caenorhabditis elegans*. In that study, GFP was placed under the transcriptional control of the *mec-7* promoter, which is activated in a small number of *C. elegans* neurons. GFP was nontoxic to cells, and its expression did not appear to interfere with cell growth and/or function. In addition, the green fluorescence did not appear to photobleach when viewed with fluorescein filter sets. These results suggested that GFP might be a powerful tool to examine changes in gene expression in living tissue.

Subsequently, GFP has been used in *Drosophila* to monitor the subcellular distribution of the exuperantia protein (exu) (6). In those studies, GFP was expressed as an in-frame fusion with the exu protein (encoded by *exu*) under the transcriptional control of its own promoter. The exu–GFP fusion protein was found to be expressed in the same pattern as the native exu protein. These results demonstrated the potential of GFP as a vital marker in *Drosophila*. However, the fact that GFP was produced as a exu–GFP fusion protein that was only expressed in adult ovaries precludes its use as a general marker/reporter gene. In theory, the exu–GFP fusion protein could be targeted to other tissues using various promoters, but the effects of ectopically expressing an exu–GFP fusion protein in other cell types are uncertain. Alternatively, additional fusion proteins could be generated with GFP, but this approach would be cumbersome, and the fusion products may be unstable, inactive, or nonfluorescent.

To make GFP more generally useful, we have utilized the GAL4 enhancer trap technique developed by Brand and Perrimon (7) to target expression of GFP. Here we show that GFP can be used as a vital marker of gene expression in a variety of living cell types at various developmental stages. We also demonstrate that GFP need not be expressed as a fusion protein in *Drosophila* but can be utilized directly as a reporter gene, much like *lacZ*. Finally, we show that GFP can be used to detect dynamic changes in gene expression in living tissue. Taken together, our results indicate that GFP will prove to be a powerful tool for viewing developmental changes within a living organism.

MATERIALS AND METHODS

Enhancer Detection Screen. GAL4-expressing enhancer trap lines were generated by mobilizing a single X chromosome-linked GAL4 *P*-element insertion (pGawB) as described (7). Four hundred crosses were set up to look for new insertion sites. In this study, only autosomal insertions were examined. Sixty-eight GAL4 insertion lines were obtained and balanced using standard genetic methods. Each of these was crossed to either Bg41-2 or Bg4-2-46 upstream activating sequence (UAS)–lacZ reporter lines, and β-galactosidase staining patterns were determined in embryos, imaginal discs, and ovaries. Out of 68 lines, 54 lines produced a detectable staining pattern. Lines that produced an interesting GAL4 expression pattern based on β-galactosidase staining were then crossed to a UAS–GFP line. The UAS–GFP line used in these experiments, GFPB1, contains a homozygous viable insertion on chromosome 3.

UAS–GFP Construct. The GFP coding region was isolated from plasmid TU#65, which contains the GFP cDNA in pBS(+) (Stratagene) as a *Kpn* I–*Eco*RI fragment and subcloned into pUAST. pUAST is a *P*-element vector based on pCaSpeR3 containing five optimal GAL4 binding sequences followed by a multiple cloning site (7).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: GFP, green fluorescent protein; UAS, upstream activating sequence; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside.

*E.Y. and K.G. contributed equally to the research presented in this paper.
†To whom reprint requests should be addressed.

Transformations. Transgenic flies carrying the UAS–GFP construct were generated by injecting pUAS–GFP DNA at a concentration of 400 μg/ml with the helper plasmid pπ25.7wc at a concentration of 100 μg/ml into embryos of the w^1 strain (8, 9) using standard methods (10). A total of eight different lines were generated with the UAS–GFP insertion on X chromosome, chromosome 2, or chromosome 3.

Detection of *lacZ* by Antibody Immunocytochemistry and 5-Bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal) Staining. The procedures used to collect embryos, remove the vitelline membrane, and stain whole-mount embryos with antibodies are described (11). The primary antibody, an IgG fraction rabbit anti-β-galactosidase from Cappel, was used at a concentration of 1:4000. The secondary antibody, a horseradish peroxidase-conjugated goat anti-rabbit IgG (Bio-Rad), was used at 1:200. Imaginal discs were dissected in chilled phosphate-buffered saline (PBS) and fixed for 20 min in 0.75% glutaraldehyde. After being washed once with PBS/0.1% Triton X-100, cells were stained as described (12) and mounted in 70% glycerol. Ovaries were dissected and fixed in 1% glutaraldehyde for 20 min. Fixative was removed, and the ovaries were stained as described (13). After being washed in PBS/0.1% Triton X-100, they were mounted in 98% glycerol. Photographs of β-galactosidase histochemistry and immunocytochemistry were taken on a Nikon Optiphot 2 microscope with Nomarski optics, using a Nikon FX-35 camera.

Visualization of GFP Using Confocal Microscopy. Confocal images of GFP expression were taken on either a Leica DM IRB inverted laser confocal microscope using a standard fluorescein isothiocyanate filter providing excitation at 490 nm and emission at 527 nm or a Bio-Rad model MRC600 using a standard fluorescein isothiocyanate filter on a Nikon Optiphot 2 microscope. In all cases, image files were processed using a computer-based graphic system (Corel 4.0) where they were arranged and annotated. Images acquired of imaginal discs, the larval nervous system, and developing egg chambers were not further processed. All other images were processed to adjust the brightness and contrast of the image using Corel PHOTOPAINT. Imaginal discs from F_1 larvae resulting from a GAL4 line/UAS–GFP cross were dissected in distilled water and mounted immediately in 70% (vol/vol) glycerol/30% 0.1 M Tris (pH = 9). Larval central nervous systems were dissected in Schneiders medium according to standard procedures (14). F_1 embryos were dechorionated with 3% sodium hypochlorite, rinsed with distilled water, and mounted in 70% glycerol/30% 0.1 M Tris (pH = 9.0). Ovaries were dissected and mounted in PBS from 2-day-old F_1 virgin females. No fixatives were used in any of these preparations. Images of GFP expression in developing ovaries were derived as follows. Ovaries were dissected in Schneider's medium/10% fetal calf serum. Stage-8 egg chambers were dissected out of the epithelial sheath overlaying ovarioles and transferred to a microscope slide with medium. An artificial well was created on the slide using stacked slips of paper covered with vacuum grease to hold enough medium to bathe the sample and to support a coverslip. A *z*-series of confocal images was then obtained every hour for a total of 4 hr.

RESULTS

To determine whether GFP could be used as a vital marker/reporter in a variety of tissues during *Drosophila* development we generated transgenic lines containing the GFP cDNA under the transcriptional control of the yeast UAS. Three independent lines were tested by crossing them to several GAL4 enhancer trap lines that we had generated according to the protocol outlined in Brand and Perrimon (7). No apparent difference was observed in the ability to detect GFP from any of the three lines tested. The results presented here were obtained with the GFP-B1 line, which is a homozygous viable insertion of the UAS–GFP transgene on chromosome 3.

Expression of GFP was first examined by crossing the UAS–GFP-B1 line to three GAL4 enhancer trap lines that are expressed within adult ovaries. In all cases, appropriate expression of the UAS–GFP transgene was confirmed by comparing the results obtained with histochemical results of parallel crosses of the GAL4 lines with a UAS–lacZ line (Fig. 1 *A*, *C*, and *E*). The GAL4 lines used in these experiments targeted expression of GFP to posterior follicle cells (Fig. 1*B*), stalk cells (Fig. 1*D*), and nurse-cell-associated follicle cells (Fig. 1*F*). Expression of GFP could be detected in both freshly dissected ovaries and in fixed tissue (data not shown). By using the GFP marker in combination with confocal microscopy, we obtained greater resolution of the expression pattern than observed using X-Gal staining. The apparent reduction in the number of cells that express GFP in the posterior and nurse-cell-associated follicle cells is due to the optical sectioning of the confocal microscope (Fig. 1 *B* and *F*). Thus, GFP can be expressed and detected in ovaries not only as a fusion protein (6) but as a reporter gene as well. As previously noted (6), we could also detect minimal levels of autofluorescence within late egg chambers. However, using a barrier filter with a wavelength cut-off of 580 nm, we could distinguish between GFP and autofluorescence: GFP emits light maximally at 509 nm and is not detectable under these conditions, whereas autofluorescence can still be observed.

We also examined whether GFP could be detected in larval tissues by crossing the UAS–GFP-B1 line to several GAL4

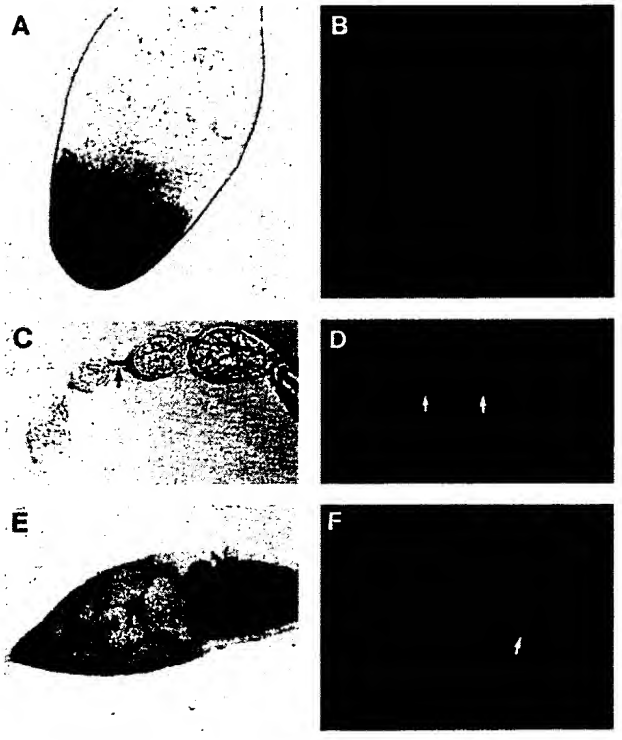


FIG. 1. Detection of GFP during oogenesis. Three GAL4 enhancer trap lines were used to detect expression of GFP during oogenesis. Expression of GFP was confirmed by crossing each GAL4 line to a UAS–lacZ line and staining ovaries with X-Gal. *A*, *C*, and *E* represent X-Gal staining; *B*, *D*, and *F* represent the same GAL4 lines examined for GFP expression. (*A* and *B*) GAL4 line A62, which directs expression of β-galactosidase and GFP in posterior follicle cells. (*C* and *D*) GAL4 line A39, which directs expression of β-galactosidase and GFP to stalk cells (see arrows). (*E* and *F*) GAL4 line A90, which directs expression of β-galactosidase and GFP to nurse-cell-associated follicle cells (see arrows). (*A–D*, ×94; *E*, ×100.)

enhancer trap lines that target GFP expression to specific cells in imaginal discs and the larval nervous system (Fig. 2). Illustrated are examples where GFP expression is targeted to the optic lobe (Fig. 2*B*), the eye disc and optic lobe (Fig. 2*D*), and a wing imaginal disc (Fig. 2*F*). Interestingly, when GFP was expressed in optic lobe neurons, the protein was not restricted to cell bodies within the larval nervous system but could also be detected in their processes. As seen in Fig. 2*D*, GFP clearly marked photoreceptor axons as they projected from the eye disc to innervate the optic lobe.

To further examine whether GFP could be used to identify nerve terminals, we used GAL4 enhancer trap lines to target expression of GFP to motoneurons within the larval nervous system (Fig. 3). These experiments clearly demonstrated that GFP could be detected both in neuronal cell bodies and in the processes immediately extending from the cell bodies (Fig. 3*A*). Similar to that observed in *C. elegans* (5), GFP could also be detected within nerve terminals at the point where they innervated specific muscles. For example, in a nerve terminal that innervates muscle 12 of the larval abdomen (Fig. 3*B*), GFP clearly outlines both the preterminal region and synaptic boutons (Fig. 3*B*). The only significant photobleaching observed was within the nerve terminals and seen only after prolonged exposure to the laser beam. However, fluorescence was recovered after a brief rest period in the absence of the laser beam. In the larval tissues examined, no autofluorescence was observed, and it was necessary to artificially increase the background by adjusting the baseline fluorescence using the black level control. This adjustment increased the total brightness of the image by a constant and allowed us to visualize the underlying structures for photography.

These experiments clearly demonstrate that GFP can be used to detect gene expression in a variety of freshly dissected tissues in *Drosophila* without any requirement for fixation or additional substrates. To determine whether GFP can also be used in intact living animals we examined GFP expression in embryos and larva (Fig. 4). Embryos were dechorionated, mounted in halocarbon oil on a glass slide, and viewed by confocal microscopy. Exposure of the embryo to the laser beam for the short periods of time required to obtain an image

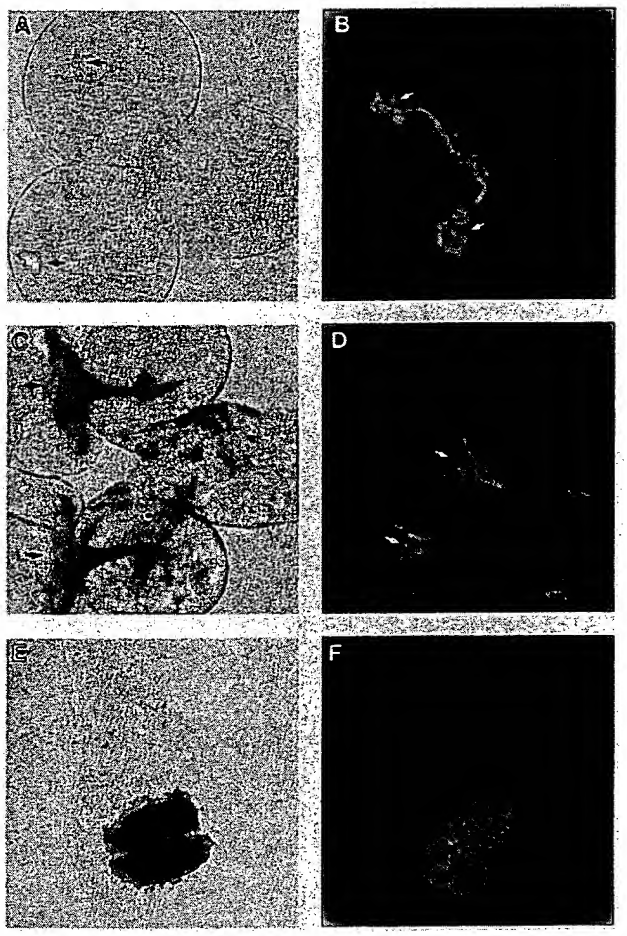


FIG. 2. Detection of GFP in larval imaginal discs and central nervous system. Three GAL4 enhancer trap lines were used to detect expression of GFP in larval imaginal discs and the central nervous system. To confirm that GFP was expressed in the appropriate pattern, each GAL4 line was crossed to a UAS–lacZ line and examined by X-Gal staining. *A*, *C*, and *E* represent the X-Gal staining pattern for each GAL4 line; *B*, *D*, and *F* represent the GFP expression pattern. (*A* and *B*) GAL4 line A95, which directs expression of β-galactosidase and GFP to neuronal cell bodies within the optic lobe and to their processes that extend from one lobe to the other (see arrows). The optic lobes are located at left and the ventral ganglia are located at right. In this preparation, the eye discs have been removed. (*C* and *D*) GAL4 line B41, which directs expression of β-galactosidase and GFP to photoreceptor neurons within the eye disc and to their processes, which innervate deep within the optic lobe. Arrows indicate location of the photoreceptor cells within the eye disc. (*E* and *F*) GAL4 line C5, which directs expression of β-galactosidase and GFP to the region of the wing imaginal disc, which will give rise to the wing blade. (*A–F*, ×110.)

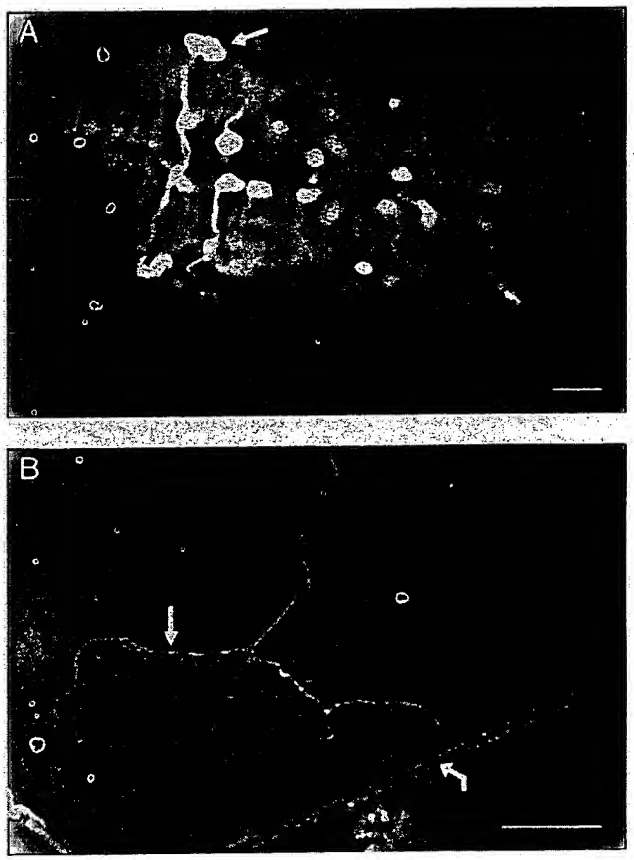


FIG. 3. Detection of GFP in motoneuron cell bodies and nerve terminals. The GAL4 enhancer trap line D42 was used to target expression of GFP to motoneurons within a living third-instar larvae. (*A*) Expression of GFP can be detected within motoneuron cell bodies and in the processes that immediately extend from them. The arrow points to a specific motoneuron within the larval CNS. (*B*) Expression of GFP can also be detected at the nerve terminal and within synaptic boutons (straight arrow). Autofluorescence can also be detected within the trachae (bent arrow). (Bar = 50 μm.)

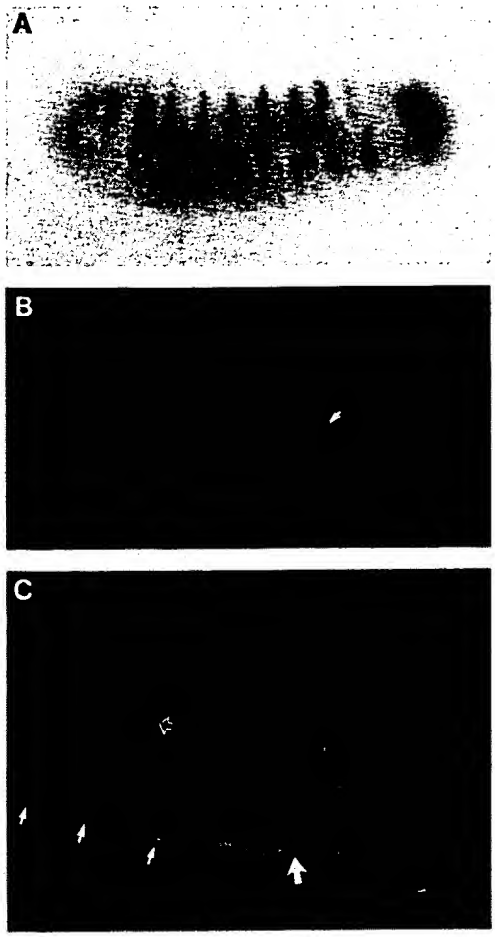


FIG. 4. Detection of GFP in living embryos and larvae. Expression of GFP can also be detected in embryos or larvae without any prior dissection or fixation. (*A*) Anti-β-galactosidase staining of embryos expressing β-galactosidase driven from the GAL4 line C41. Expression can be detected in most, but not all, of the cells which comprise the peripheral nervous system of the embryo. The arrow points to a group of cells within the lateral cluster of the peripheral nervous system. (*B*) Expression of GFP driven by the GAL4 line C41 can be detected within the peripheral nervous system of live embryos. (*C*) Expression of GFP driven by the GAL4 line C38 can be detected within the larval salivary glands (large filled arrow) and pair of cells along the body wall within each segment (small filled arrows). Autofluorescence from the gut is indicated by an open arrow. (*A–C*, ×55.)

(<1 min) did not alter its viability, and embryos were observed to hatch into larvae (data not shown). Expression of GFP is shown for one GAL4 line, C41, which is expressed in much of the embryonic peripheral nervous system (Fig. 4*B*). Appropriate expression of GFP was confirmed by immunostaining embryos with an anti-β-galactosidase antibody (Fig. 4*A*). While all of the peripheral neurons cannot be observed in a single confocal optical section, the overall pattern is maintained. Similar observations can also be made in intact larvae. Fig. 4*C* shows a GAL4 enhancer trap line that targets GFP to both salivary glands and pairs of cells that are distributed segmentally along the larval body wall. In larvae, as in embryos, development was not affected by exposure to the laser beam. In contrast to larval imaginal discs and nervous system, whole embryos and larvae have significant levels of autofluorescence due to the yolk and gut, respectively. These levels, however, could be resolved by using appropriate barrier filters as described above.

To determine whether GFP could be used to detect dynamic changes in gene expression within living tissues in real time, we

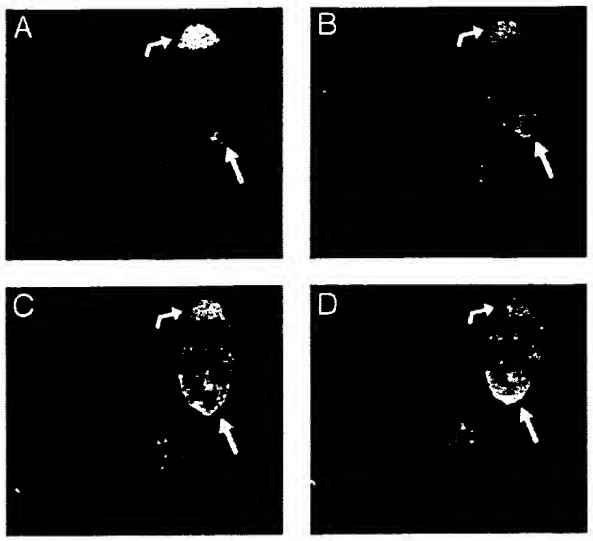


FIG. 5. Changes in GFP expression can be detected in cultured egg chambers. Expression of GFP was monitored in cultured stage-8 egg chambers during a 4-hr period. A *z*-series of confocal images was obtained every hour for the entire 4-hr period. Images represent a particular focal plane from each *z*-series. Expression of GFP in egg chambers is directed by GAL4 line A90, which targets expression to nurse-cell-associated follicle cells. (*A*) Detection of GFP after 1 hr in culture. Little to no expression of GFP can be detected within follicle cells (straight arrow). Low levels of autofluorescence, however, can be observed within the oocyte (bent arrow). (*B*) GFP expression begins to be detected after 2 hr in culture in follicle cells at the anterior end of the stage-8 egg chamber (straight arrow). In contrast, the level of autofluorescence from the oocyte decreases (bent arrow), and no changes are observed in earlier egg chambers. (*C*) After 3 hr, levels of GFP expression within nurse-cell-associated follicle cells increase (straight arrow). (*D*) After 4 hr, GFP expression can be detected in all nurse-cell-associated follicle cells and is particularly high in follicle cells at the anterior tip of the egg chamber (straight arrow). No changes in fluorescence are detected in earlier egg chambers, and the autofluorescence in the oocyte remained low (bent arrow). (*A–D*, ×90.)

performed time-lapse confocal microscopy of developing egg chambers (Fig. 5). These experiments were done by using a GAL4 enhancer trap line, A90, which targets expression of UAS–GFP to nurse-cell-associated follicle cells (Fig. 1*C*). Egg chambers that were dissected and cultured in Schneiders medium over a period of 4 hr are shown. Initially, no GFP could be detected within stage-8 egg chambers (Fig. 5*A*). However, by 1 hr, GFP expression was observed at the anterior end of the stage-8 egg chamber, and this expression increased steadily over time (Fig. 5 *B–D*). Expression of GFP is specific to nurse-cell-associated follicle cells and restricted to stage-8 egg chambers. No GFP could be detected in stage-2 to -7 egg chambers. These results clearly show that GFP can be used as a reporter to monitor activation of gene expression in living tissue over time. Whether GFP can also be used to monitor cessation of gene expression remains to be determined and will depend on the stability of GFP in various cell types during development.

DISCUSSION

The ability to study development as it occurs within an organism relies on the availability of techniques that can detect changes in gene expression within specific cells or tissues during cell movements and migrations. We have used the GFP from the jellyfish, *A. victoria*, as a viable marker in *Drosophila* to observe such changes within living tissues. GFP was expressed as a nonfused protein under the transcriptional control of a yeast UAS and targeted to specific cell types during

development by crossing to a variety of GAL4-expressing enhancer trap lines. These studies clearly show that GFP can be used as a reporter gene, much like bacterial *lacZ*, to detect expression of specific genes in a variety of cell types during *Drosophila* development. However, in contrast to *lacZ*, GFP can be visualized in live tissue without fixation or addition of specific substrates and often without any dissection. This result permits the monitoring of gene expression within a living organism over time.

The ability to use GFP as a vital marker/reporter in *Drosophila* suggests a number of other interesting applications. For example, development of an enhancer trap vector system based on GFP suggests the possibility of bulk screening, whereby specific expression patterns could be detected in embryos in the F_1 generation. Because GFP can function as a reporter gene, it should be able to replace *lacZ* in other assays, such as promoter mapping. In addition, it may be possible to use GFP to sort pure populations of live cells using a fluorescence-activated cell sorter as originally described by Krasnow *et al.* (15) using a fluorogenic β-galactosidase substrate. GFP could also be recombined onto various balancer chromosomes to allow for rapid identification of embryos/larvae containing the specific balancer, much like *lacZ*-marked balancers that are currently available. GFP balancers would permit identification and selection of homozygous mutants based on the absence of GFP expression. *P*-element vectors could also be generated using GFP as a reporter gene and used to characterize various mutant phenotypes arising from *P*-element insertional mutagenesis. Because GFP is expressed not only in cell bodies but also in processes, this may be particularly useful in identifying axon guidance or pathfinding mutants. Until recently, these mutants were difficult to identify, as β-galactosidase fails to readily diffuse into axons. Alternative approaches in which β-galactosidase is expressed as a kinesin–lacZ (16) or Tau–lacZ fusion protein (17) are more efficient at detecting axonal processes but still require fixation and the addition of specific substrates or antibodies to detect expression, which only provides a static image of the axonal process. The expression of GFP in nerve terminals may also be useful for studying synaptogenesis. For example, GAL4 lines that target expression of GFP to photoreceptor axons could be used to examine their ability to form appropriate synapses within the optic lobe. Expression of GFP in nerve terminals could also be used to identify specific neurons and/or synaptic boutons, which could then be analyzed electrophysiologically.

Finally, we have shown that GFP can be used to detect changes in gene expression in living tissue. This result suggests that GFP could be used in fate mapping or lineage analysis experiments. GFP could also be used to monitor changes in cell migrations or cell shape such as occur during germ-band extension (18) and the migration of pole cells (19), follicle cells (13), and tracheal cells (20, 21) during development. The identification of mutations that affect the pattern of specific cell migrations combined with the ability to visualize the cells in living tissues using GFP should provide insight into the mechanisms that control cell movements as they occur within the organism.

We thank Drs. A. Brand and N. Perrimon for kindly providing us with the pUAST vector, the GAL4–lethal/FM7 jumpstarter and UAS–lacZ strains and Dr. M. Chalfie for the GFP cDNA. We thank L. Zhou for providing us with excellent technical help during the course of these experiments and B. Stewart for help with larval central nervous system preparations. We are indebted to J. Wadia, J. Georgiou, and Drs. W. G. Tatton and M. Charlton for excellent advice and help with all aspects of confocal microscopy. The Bio-Rad MRC600 confocal microscope facility was provided by the Neuroscience Network. We are most grateful to Drs. W. S. Trimble, H. Krause, and J. Culotti for critical comments on the manuscript. This work was supported by the National Science and Engineering Council of Canada, the Medical Research Council of Canada, and the Amyotrophic Lateral Sclerosis Society of Canada. E.Y. is partially supported by a University of Toronto Open Scholarship. G.L.B. is supported by the Alzheimers Association of Ontario.

1. Silhavy, T. J. & Beckwith, J. R. (1985) *Microbiol. Rev.* **49**, 398–418.
2. O'Kane, C. J. & Gehring, W. J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 9123–9127.
3. Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K. & Gehring, W. J. (1989) *Genes Dev.* **3**, 1288–1300.
4. Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., Ackerman, L., Caretto, R., Uemura, T., Grell, E., Jan, L. Y. & Jan, Y. N. (1989) *Genes Dev.* **2**, 1273–1287.
5. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. (1994) *Science* **263**, 802–805.
6. Wang, S. & Hazelrigg, T. (1994) *Nature (London)* **369**, 400–403.
7. Brand, A. H. & Perrimon, N. (1993) *Development (Cambridge, U.K.)* **118**, 401–415.
8. Rubin, G. M. & Spradling, A. C. (1982) *Science* **218**, 348–353.
9. Karess, R. E. & Rubin, G. M. (1984) *Cell* **38**, 135–146.
10. Spradling, A. C. (1986) in *Drosophila: A Practical Approach*, ed. Roberts, D. B. (IRL, Oxford), pp. 175–197.
11. Bodmer, R. & Jan, Y. N. (1987) *Roux Arch. Dev. Biol.* **196**, 69–77.
12. Wagner-Bernholz, J. T., Wilson, C., Gibson, G., Schuh, R. & Gehring, W. J. (1991) *Genes Dev.* **5**, 2467–2480.
13. Montell, D. J., Rorth, P. & Spradling, A. C. (1992) *Cell* **71**, 51–62.
14. Budnick, V., Zhong, Y. & Wu, C.-F. (1990) *J. Neurosci.* **10**, 3754–3768.
15. Krasnow, M. A., Cumberledge, S., Manning, G., Herzenberg, L. A. & Nolan, G. P. (1991) *Science* **251**, 81–85.
16. Giniger, E., Wells, W., Jan, L. Y. & Jan, Y. N. (1993) *Roux Arch. Dev. Biol.* **202**, 112–122.
17. Callahan, C. A. & Thomas, J. B. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5972–5976.
18. Irvine, K. D. & Wieschaus, E. (1994) *Development (Cambridge, U.K.)* **120**, 827–841.
19. Jaglarz, M. K. & Howard, K. R. (1994) *Development (Cambridge, U.K.)* **120**, 83–89.
20. Klambt, C., Glazer, L. & Shilo, B.-Z. (1992) *Genes Dev.* **6**, 1668–1678.
21. Reichman-Fried, M., Dickson, B., Hafen, E. & Shilo, B.-Z. (1994) *Genes Dev.* **8**, 428–439.

THIS PAGE BLANK (USPTO)

Production of Protein Pharmaceuticals in Transgenic Plants

Copy to
Winstone
2/7/29

Philip C. Lyons
Plant Biotechnology Program, Texas A & M University
Institute of Biosciences and Technology, Houston, TX 77035, USA

Gregory D. May, Hugh S. Mason *and* ***Charles J. Arntzen***
Plants and Human Health Program, Boyce Thompson Institute for Plant Research
Tower Road, Ithaca, NY 14853-1801, USA

Introduction

Genetic transformation techniques are now applied routinely to a large number of plant species. Through this technology, numerous genes that confer agronomically important traits such as pest resistance and herbicide tolerance have been introduced into plants. The successes of modern plant biotechnology exemplified by these achievements in crop improvement have created considerable interest in further exploiting the remarkable biosynthetic capacity of plants by developing transgenic plants that will produce valuable new products. With a number of recent demonstrations that transgenic plants are capable of producing functional foreign proteins and peptides of known or potential pharmaceutical importance (Table I), new and exciting possibilities have been created in the ancient science of plant medicinal chemistry.

Biologically active peptides and proteins have many potential pharmaceutical applications, including use as vaccines, immunomodulators, growth factors, hormones, blood proteins and enzymes. Efforts to produce these compounds economically and in adequate quantities have become increasingly reliant upon the use of various prokaryotic and eukaryotic cell-culture expression systems. In this article, we will consider the use of transgenic plants as alternative eukaryotic expression systems for the production of recombinant protein pharmaceuticals, and the advantages of plants compared to other systems.

Expression in Transgenic Plants

For many reasons, transgenic plants are a feasible, and in some cases preferable, eukaryotic expression system for the production of valuable pharmaceuticals. The number of

Table 1

Peptides or proteins with known or potential pharmaceutical applications that have been expressed in transgenic plants.

Protein or Peptide	Application
Hepatitis B surface antigen	*Vaccine*
Norwalk virus capsid protein	*Vaccine*
Foot-and-mouth disease virus	*Vaccine*
Human rhinovirus 14	*Vaccine*
Human immunodeficiency virus	*Vaccine*
S. mutans surface protein	*Vaccine*
E. coli enterotoxin, B subunit	*Vaccine*
Malarial circumsporozoite epitopes	*Vaccine*
Mouse ZP3 protein epitope	*Vaccine*
Mouse catalytic antibody 6D4	*Antibody*
Mouse mAB Guy's 13	*Secretory Antibody*
Mouse Mab B 1-8	*Antibody*
Anti-phytochrome Fv protein	*Antibody*
Anti-substance P	*Antibody*
Human serum albumin	*Serum Protein*
Human protein C	*Serum Protein*
α-trichosanthin	*Cytotoxin*
Ricin	*Cytotoxin*
Human epidermal growth factor	*Neuropeptide*

plant species amenable to genetic transformation is now quite large [1,2]. These plants have the capacity to express foreign genes from a wide range of sources, including viruses, bacteria, fungi, insects, animals and other plants. Plants also are capable of high levels of protein expression; foreign protein concentrations as great as 30% of the total soluble protein have been reported in transgenic plants [3], and expression levels in excess of 1% are often attainable [4-8]. With rapid advances being made in the manipulation of foreign protein expression through the development of novel vectors and improved understanding of protein folding, assembly and processing in plants, one can expect that more consistent and higher levels of expression for a wider range of functional proteins and peptides will be attained.

Biologically active peptides and proteins have many potential pharmaceutical applications.

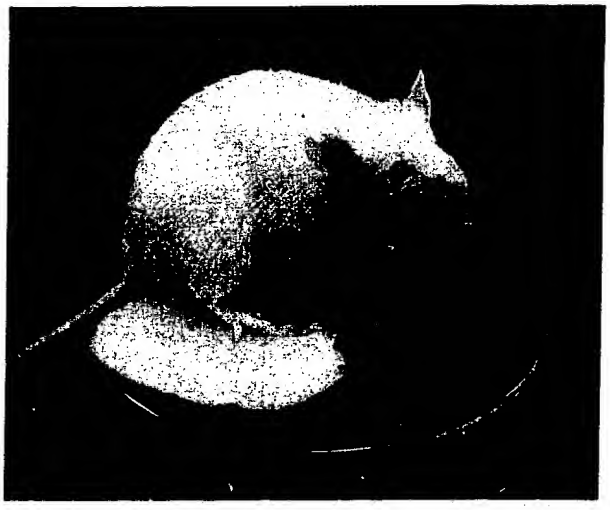
With respect to protein production, plants have an advantage of low cost involved in growing large amounts of biomass. With many of the difficulties traditionally encountered in achieving adequately purified protein from plants now overcome through advancements in protein purification techniques, it should now be possible to take greater advantage of plants as efficient sources of recombinant proteins. Furthermore, should it prove feasible to utilize transgenic plants as an edible source for the oral delivery of recombinant pharmaceuticals — as some evidence already suggests [9,10] — the need for costly purification procedures would in such instances be entirely eliminated.

Foreign proteins can be expressed in plants either transiently or as stably inherited traits [1,2,10]. Both approaches have been applied to the production of protein pharmaceuticals. Transient expression of pharmaceutically valuable proteins has been achieved most often by infection with a plant virus into which the gene of interest has been introduced through recombination [10]. Various modifications of this approach have been used with several different plant viruses. Infection with the genetically modified virus leads to the production of a fusion protein, typically consisting of the virus coat protein fused to a small foreign peptide sequence. Conversely, stable transformation is achieved by incorporation into the plant genome, by *Agrobacterium*-mediated transformation [1,2,10], of the foreign gene(s) encoding the protein of interest. Incorporation into the genome is the basis for the stability and heritability of traits introduced by this method.

Recombinant Vaccines in Transgenic Plants

Vaccines against viruses. A substantial number of candidate recombinant vaccines have been produced in plants through either stable or transient gene expression. Human virus diseases have been the most frequent target for plant-recombinant vaccines. The first of these was the hepatitis B surface antigen (HBsAg) [11]. Plant-derived recombinant HBsAg (rHBsAg) was expressed in tobacco plants (an important model plant species widely used for transgenic research). Although the expression levels were low (~0.01% of total soluble protein) it was shown in these experiments that the plant derived rHBsAg assembles into virus-like particles (VLPs). These particles were similar in size and other physical properties to yeast rHBsAg, which is the source of commercial hepatitis B vaccine (Recombivax®; Merck, Sharpe & Dohme). When a partially purified preparation of the tobacco-derived rHBsAg was parenterally injected into mice, it resulted in an antibody response that mimicked the response obtained with Recombivax [12]. In addition, lymph node T-cells isolated from mice, which were primed with tobacco-derived rHBsAg, could be stimulated to proliferate *in vitro* by both tobacco-derived and yeast-derived rHBsAg. These experiments demonstrate the close antigenic relatedness of the plant-derived recombinant protein to a known, highly effective recombinant vaccine produced in another eukaryotic expression system.

Another potential virus vaccine to be expressed in transgenic plants is the Norwalk virus capsid protein (NVCP) [13]. Norwalk virus causes epidemic acute gastroenteritis in humans. Plant-derived recombinant NVCP (rNVCP) has been expressed stably in both tobacco leaves and potato tubers, and like plant-derived rHBsAg, it self-assembled into VLPs. Expression and self-assembly of the capsid protein had first been achieved in recombinant baculovirus-infected insect cells [14] and this material has shown promise as an oral vaccine. In plant cells, the accumulation of rNVCP was nearly 0.3% of total protein; for potatoes, this resulted in a yield of about 20 µg rNVCP per gram of tuber weight. Feeding transgenic tubers expressing rNVCP to mice caused both humoral and mucosal anti-Norwalk virus antibodies to be produced. These results demonstrated that a plant-derived recombinant subunit antigen causes oral immunization when consumed as a food



Transgenic potatoes expressing a cholera vaccine have caused an oral immunization in mice fed the potatoes.

mutans, the principal cause of dental caries in humans. In this instance, four distinct components (a heavy and light chain, and two additional proteins, one of which is added during secretion) were expressed in a single transgenic plant. Furthermore, part of the heavy chain constant regions were replaced with domains from an IgA heavy chain since the secretory IgA form of antibody has proven more effective in passive immunotherapy against *S. mutans*. Remarkably, despite the heavy chain modification and the complexity of the recombinant molecule, all four proteins assembled into a functional antibody that recognized the native antigen from *S. mutans* and caused aggregation of cells of this bacterium. It was reported that efforts are now in progress to formulate the plant-derived recombinant antibody into toothpaste to determine if this is an effective means to prevent tooth decay.

A number of proteins present in human serum are of vital importance to medicine.

Other approaches have been used to produce antibodies in plants as well. In one instance [26], chimeric genes, containing the light or heavy chain of mAb B 1-8 fused with a plant signal peptide, were placed under the control of separate promoters in a single plasmid. Tobacco plants transformed with the chimeric genes produced both the light and heavy chains, which assembled into immunologically detectable antibodies that bound to the appropriate hapten. The level of expression was not determined in this study.

Another approach, first developed for antibody expression in *E. coli* [27], involves expression of single-chain Fv antibodies in which the light and heavy chain variable domains of an immunoglobulin are joined together by a flexible peptide linker. The flexible linker facilitates folding and assembly of the two chains, resulting in a functional synthetic antibody fragment. This approach was used successfully to express a functional anti-phytochrome single-chain Fv protein in transgenic tobacco [28]. It also was used with tobacco to express the antibody to a coat protein of the artichoke mottled crinkle virus [29], resulting in a reduced incidence of infection and delayed symptom development in virus-inoculated transgenic tobacco plants. Expression levels of the single-chain Fv antibodies in each instance was about 0.1% of total soluble leaf protein.

One additional approach used for antibody expression in plants involved the construction of a general purpose vector with a multiple cloning site that allows the insertion of a heavy chain variable (VH) domain [30]. VH domains expressed in *E. coli* have been shown to fold correctly and often retain antigen-binding activity. Using this vector, a "single-domain antibody" consisting of the VH domain of antibody to substance P (a neuropeptide) was expressed in transgenic tobacco, where it accumulated to approximately 1% of total soluble protein.

Serum Proteins in Transgenic Plants

A number of proteins (including coagulation and anti-coagulation factors) present in human serum are of vital importance to medicine. Due to the nature of donated plasma and the necessity for a sufficient supply of highly purified blood products, alternative production systems are required. Recombinant human serum albumin (rHSA) that is indistinguishable from the authentic human protein has been expressed in transgenic tobacco and potato [31]. The significance of this particular accomplishment rests not only in the demonstration that a valuable protein is produced in transgenic plants, but also that it was possible to achieve proper processing by fusion of rHSA to a plant presequence, resulting in cleavage and secretion of the correct protein. This achievement was particularly relevant in view of the substantial difficulties that have been encountered in other systems used to express recombinant HSA. The level of expression in transgenic potato plants was 0.02% of total soluble leaf protein.

Human serum protein C, a highly modified and glycosylated serine protease zymogen that requires proteolytic processing, also has been produced in transgenic tobacco plants (D.L. Weissenborn, Virginia Polytechnic Institute and State University, unpublished data). In this research, tobacco plants expressed both single chain and heavy chain forms of the protein, suggesting that the tobacco-derived recombinant protein C undergoes proper cleavage as well.

Recombinant Toxins in Transgenic Plants

Many toxins are proteins, usually derived from bacteria or plants, which kill cells by interfering with metabolism, often by the inhibition of protein synthesis. Due to the extreme cytotoxicity of many of these proteins, they have been the subject of intense investigation as tumor controlling agents. Most toxins exhibit very little site specificity, however, and in order to be used therapeutically, require modification of the binding domains so that the toxin will be preferentially directed to the appropriate target [32]. The modification usually involves inactivation or removal of the binding domain and expression of the toxin as a fusion protein with a vector that can deliver the toxin to the preferred site of action.

There have been two recombinant toxins expressed in transgenic plants to date [4,7], both of which are plant-derived eukaryotic ribosome-inactivating proteins (RIPs). The first of these, α-trichosanthin, a 27 kDa protein from *Trichosanthes kirilowii*, inhibits the replication of HIV in acutely infected $CD4^+$ lymphoid cells and in chronically infected macrophages. Although α-trichosanthin had been expressed previously in *E. coli*, the amount recovered was low, less than 0.01% of total cellular protein. Transient expression of this protein in tobacco plants using recombinant TMV as the vector resulted in accumulation of α-trichosanthin in leaves to at least 2% of total soluble protein 141.

Ricin is an RIP from castor bean that has therapeutic potential for the treatment of AIDS and cancer. Development of recombinant toxins utilizing ricin also has been hampered by low expression levels, as well as difficulties in achieving proper folding and processing in various expression systems.

However, stable expression of active, processed recombinant ricin was achieved in transgenic tobacco [7]. As with the transiently expressed α-trichosanthin, levels of the stably expressed recombinant ricin reached approximately 2% of the total soluble protein. The results obtained with these two toxins, together with the successful expression of fusion proteins described elsewhere in this article, suggest that transgenic plants may be quite useful for the production of site-selective toxins requiring fusion to a targeting protein.

Bioactive Peptides in Transgenic Plants

In addition to the vaccine epitopes discussed above, other small bioactive peptides with various potential pharmaceutical applications have been expressed in transgenic plants. Stable expressson of a chemically synthesized gene for human epidermal growth factor (hEGF), a small mitogenic peptide which stimulates *in vitro* proliferation of animal cells, was achieved in transgenic tobacco [33]. Although native hEGF is processed proteolytically in human cells from a larger precursor, the synthetic gene encoded only for the active peptide portion. Nevertheless, incorporation of this synthetic gene into the plant genome resulted in expression of a peptide that reacted with hEGF specific antibody. However, for plants to be feasible as a source of hEGF, expression levels will need to be increased, since the highest content of hEGF measured in transgenic plants was only 0.001% of total soluble proteins.

Leu-enkephalin, a pentapeptide from brain that exhibits opiate activity, has been produced using two different approaches. It was first expressed transiently in tobacco protoplasts as a fusion protein with the TMV coat protein, where it represented the major protein in the cells [34]. Subsequently, it was expressed as a fusion protein with a 2S albumin seed storage protein [35] in whole transgenic *Arabidopsis thaliana* and *Brassica napus* (oilseed rape) plants. Because 2S albumins represent up to 60% of total seed protein, depending upon the plant species, expression of foreign proteins in this manner should greatly facilitate protein purification. Using a 2S albumin-rich fraction from either species, it was possible to obtain highly purified peptide by tryptic digestion and HPLC. The yield obtained from *B. napus* was estimated to be equivalent to 15-75 g of peptide per hectare. Furthermore, the authors of this study suggested that by screening individual transformants for higher expression levels—by using lines of oilseed rape with higher protein/oil ratios and by standard breeding techniques—yields of the pentapeptide could be increased substantially.

An inhibitor of angiotensin I converting enzyme, found in the tryptic hydrolysate of milk, was expressed transiently in transgenic tobacco and tomato as a fusion protein with the TMV coat protein [9]. This 12-peptide inhibitor has anti-hypertensive effects when orally administered. Yield of the fusion protein in tomato fruit was approximately 10 μg fusion protein per g of plant tissue, equivalent to approximately 0.7 μg of peptide. It was suggested that the tomato fruit could be administered orally as a dietary antihypertensive agent, since the inhibitor should be released from the fusion protein in the intestine by trypsin digestion.

Conclusion

There is now ample evidence that transgenic plants are a feasible alternative system for the production of recombinant protein pharmaceuticals. From the examples that have been discussed here, it is apparent that this capability extends to a wide variety of pharmacologically active compounds. Although levels of accumulation of the recombinant proteins considered in this discussion varied considerably, from about 0.001% of total soluble protein to greater than 2%, in many instances, the levels achieved were comparable to, or greater than, that achieved with other systems. Furthermore, it is reasonable to expect that increased levels of production for protein pharmaceuticals are attainable, particularly since accumulation of non-pharmaceutical recombinant proteins to levels as great as 30% of total soluble protein has been reported. Particularly encouraging is the degree of success that has been attained in producing active, processed forms of the expressed candidate protein pharmaceuticals. Evidence for appropriate post-translational processing steps such as folding, assembly, secretion, and proper cleavage of precursor molecules indicate that even highly complex foreign proteins can be produced in plants and are likely to be functional.

There are certain aspects of transgenic plants beyond the usual considerations of protein expression levels that would, in some instances, make their use as an expression system preferable. For example, attempts to express RIPs and other cytotoxins to high levels in other systems have encountered difficulties because of their extreme toxicity. Plants, from which many cytotoxins are derived, are in some instances considerably less sensitive than other organisms. Another aspect concerns the prospect of delivering orally administered vaccines and other pharmaceuticals via edible plant tissues expressing these bioactive compounds. The eco-

Vegetables are known sources of nutraceuticals. The next step is to produce transgenic plants with novel pharmaceutical properties.

nomics and logistics of this approach are ideally suited to developing countries, where transportation and an adequate cold-chain — necessary for most current vaccines and many other drugs — are lacking.

The use of transgenic plants is, of course, not strictly limited to protein or peptide pharmaceuticals. In the future, more genes involved in the biosynthesis of non-protein pharmaceuticals will become available. Together with increased understanding of protein expression in plants, this should lead to greater reliance in the pharmaceutical industry upon transgenic plants for the production of both protein and non-protein drugs.

References

1. C.S. Gasser and R.T. Fraley, *Sci. Amer.*, 266, 62 (1992).
2. H. Uchimaya et al., *J. Biotechnology*, 12, 1 (1989).
3. K.E. McBride et al., *Proc. Natl. Acad. Sci. USA*, 91, 7301 (1994).
4. M.H. Kumagai et al., *Proc. Natl. Acad. Sci. USA*, 90, 427 (1993).
5. K.E. McBride et al., *BioTechnology*, 13, 362 (1995).
6. G.J.H. Van Rooijen and M.M. Moloney, *BioTechnology*, 13, 72 (1995).
7. P.C. Sehnke et al., *J. Biol. Chem.*, 269, 22473 (1994).
8. R.E. Shade et al., *BioTechnology*, 12, 793 (1994).
9. H. Hamamoto et al., *BioTechnology*, 11, 930 (1993).
10. H.S. Mason and C.J. Arntzen, *Trends Biotech.*, 13, 288 (1995).
11. H.S. Mason et al., *Proc. Natl. Acad. Sci. USA*, 89, 11745 (1992).
12. Y. Thanavala et al., *Proc. Natl. Acad. Sci. USA*, 92, 3358 (1995).
13. H.S. Mason et al., *Proc. Natl Acad. Sci. USA*, (in press).
14. X Jiang et al., *Science*, 250, 1580 (1990).
15. C.J. Arntzen et al., in *Vaccines '94, Modern Approaches to New Vaccines Including Prevention of AIDS*, 339 (1994).
16. R. Usha et al., *Virology*, 197, 366 (1993).
17. C. Porta et al., *Virology*, 202, 949 (1994).
18. L. McClain et al., *AIDS Res. Human Retrovirus*, 11, 327 (1995).
19. R. Curtiss and G.A. Cardineau, World Intellectual Property Organization PCT/US89/03799 (1990).
20. J.D. Clemens et al., *J. Infect. Dis.*, 158, 372 (1988).
21. T.A. Haq et al., *Science*, 268, 714 (1995).
22. T.H. Turpen et al., *BioTechnology*, 13, 53 (1995).
23. J. Fitchen et al., *Vaccine*, 13, 1051 (1995).
24. A. Hiatt et al., *Nature*, 342, 76 (1989).
25. J.K-C. Ma et al., *Science*, 268, 716 (1995).
26. K. During et al., *Plant Mol. Biol.*, 15, 281 (1990).
27. R.E. Bird et al., *Science*, 242, 423 (1988).
28. M Owen et al., *BioTechnology*, 10, 790 (1992).
29. P. Tavladoraki et al., *Nature*, 366, 469 (1993).
30. E. Benvenuto et al., *Plant Mol. Biol.*, 17, 865 (1991).
31. P.C. Sijmons et al., *BioTechnology*, 8, 217 (1990).
32. R.J. Kreitman and I. Pastan, *Advances in Pharmacology*, 28, 193 (1994).
33. K. Higo et al., *Biosci. Biotech. Biochem.*, 9, 1477 (1993).
34. N. Takamatsu et al., *FEBS Lett.*, 269, 73 (1990).
35. J. Vandekerckhove et al., *BioTechnology*, 7, 929 (1989).

Charles J. Arntzen *is the President and CEO of the Boyce Thompson Institute for Plant Research, Inc., a not-for-profit corporation affiliated with Cornell University which is devoted to plant science. He was fomerly Dean and Deputy Chancellor for Agriculture at Texas A&M University.*

Hugh Mason's *research focus in the past three years has been the use of plants as a recombinant expression system for foreign proteins, especially candidate vaccine antigens, and the testing of transgenic plant material for immunogenic activity in animals.*

Greg May *joined the Plants and Human Health Program at the Boyce Thompson Institute for Plant Research in December 1995. His current research efforts are directed at understanding the fundamental regulatory controls of a number of complex developmental plant processes.*

Phil Lyons *was at the Alkek Institute of Biosciences and Technology at the Texas Medical Center, where he was involved in a research project on the expression of antigenic proteins in transgenic plants.*

Comments? Questions? Critiques?

Send all correspondence to:
Pharmaceutical News
c/o 820 Town Center Drive
Langhorne, PA 19047 USA

Interested in contributing a technical review?

Contact:
John Vassallo
Pharmaceutical News
c/o 820 Town Center Drive
Langhorne, PA 19047 USA
Telephone: 215-750-2642; Fax: 215-750-6343

THIS PAGE BLANK (USPTO)

Transgenic plants as vaccine production systems

Hugh S. Mason and Charles J. Arntzen

Transgenic plants that express foreign proteins with industrial or pharmaceutical value represent an economical alternative to fermentation-based production systems. Specific vaccines have been produced in plants as a result of the transient or stable expression of foreign genes. It has recently been shown that genes encoding antigens of bacterial and viral pathogens can be expressed in plants in a form in which they retain native immunogenic properties. Transgenic potato tubers expressing a bacterial antigen stimulated humoral and mucosal immune responses when they were provided as food. These results provide 'proof of concept' for the use of plants as a vehicle to produce vaccines.

The development of genetic transformation technology for plants has facilitated the study of plant gene expression, and has resulted in great progress toward the genetic design of plants with enhanced production traits (such as herbicide, insect and disease resistance). Recently, several academic and industrial laboratories have begun experimenting with transgenic plants as novel manufacturing systems. We introduced the concept of vaccine production in transgenic plants in 1992 (Ref. 1). This effort was stimulated by interest in evaluating the capacity of plants to produce different classes of proteins of pharmaceutical value, and because of the practical need for new technology for the production and delivery of inexpensive vaccines, especially in the developing world. Candidate vaccines should be a useful test system for evaluating the capacity of transgenic plants to produce pharmaceutically active proteins, because the immune system would amplify the biological response to even relatively low levels of foreign protein.

On a more applied level, the announcement of the Children's Vaccine Initiative[2] documented the need for new vaccine technology to combat infectious diseases. We hypothesized that plants could be a useful system for producing vaccines, because large amounts of antigen could be produced at a relatively low cost, using agriculture instead of sophisticated and expensive cell culture-based expression systems. In this review, we will discuss the progress that has been made by several groups in what is now an expanding area of vaccine research that utilizes transgenic plants.

Oral vaccines and mucosal immune responses

Many infectious agents colonize or invade epithelial membranes; these include bacteria and viruses that are transmitted in contaminated food or water or by sexual contact. Vaccines that are effective against these infections must stimulate the mucosal immune system to produce secretory IgA (S-IgA) at mucosal surfaces such as the gut and respiratory epithelia[3,4]. In general, a mucosal immune response is more effectively achieved by oral, rather than parenteral, antigen delivery. Several particulate antigens have proven to be effective oral immunogens, including live and killed microorganisms. By comparison with parenteral immunizations, oral immunization using subunit or soluble antigens is often inefficient at stimulating an immune response, and requires larger amounts (mg versus μg) of antigen[5].

Subunit vaccines based upon recombinant cell-culture expression systems are feasible but, for commercial-scale production, these systems require fermentation technology and stringent purification protocols so that sufficient amounts of recombinant protein can be obtained for oral delivery. Even with technological improvements, fermentation-based subunit vaccine production may be a prohibitively expensive technology for developing countries where novel oral vaccines are urgently needed. Transgenic plants that express antigens in their edible tissue might be used as an inexpensive oral-vaccine production and delivery system[6]; therefore, immunization might be possible simply through consumption of an 'edible vaccine'.

The choice of which antigens to use in the initial studies has been strongly influenced by the desire to determine if transgenic plant materials containing foreign antigens will result in oral immunization and stimulate a mucosal immune response. Because of the need in the developing world for new oral-vaccine technology against diarrheal diseases[2], antigens from enteric pathogens have been the early targets for plant-based expression. In addition, effort has focused on the

H. S. Mason and C. J. Arntzen are at the Boyce Thompson Institute for Plant Research, Tower Road, Ithaca, NY 14853-1801, USA.

© 1995 Elsevier Science Ltd

production of antigens that assemble into ordered structures, such as virus-like particles (VLP), with the hope that they will be more resistant to digestion, more likely to reach the gut-associated lymphoid tissue (GALT) and, thus, more likely to be perceived as a foreign antigen once it reaches the GALT (Refs 3,4).

The choice of plant system for testing recombinant antigen production was initially driven by convenience and the need to evaluate genetic constructs quickly. For this reason, tobacco plants were generally utilized[1,7] but, because of the high levels of toxic alkaloids in the leaves, studies on animals feeding on tobacco are not practical without substantial purification of tobacco-derived antigens. Therefore, potato has been used as the plant of choice in several of the studies reported below[6,7]. This decision was based upon the observation that mice would accept raw potato tubers in lieu of laboratory feed, and because tubers can be generated for feeding studies within a few months of the transformation process[8]. With the purpose of acquiring a delivery system for humans, we have recently developed a genetic transformation system for banana[9], and are investigating the requirements for expressing abundant proteins in banana fruit. Crops used for animal feed, such as alfalfa, grains and beans, are obvious choices for animal vaccines, although the difficulty of generating uniform transgenic samples of these species has prohibited their use in most early studies.

Systems for expressing foreign proteins in plants

Two different strategies for transgene expression (for candidate vaccine production) in plants have been evaluated (Fig. 1). These involve either: (1) stable genomic integration, with foreign DNA introduced either by *Agrobacterium* T-DNA vectors or by direct means (including microprojectile bombardment); or (2) transient expression using viral vectors. Stable expression affords the advantage of the subsequent generation of large numbers of transgenic plants, either by vegetative or sexual means, and the opportunity to introduce more than one gene for possible multi-component vaccine production. In addition, judicious choice of genetic regulatory elements allows organ and tissue-specific expression of foreign antigens. Transient expression is less easy to initiate, because the viral vector must be inoculated into individual host plants; however, greater yields of foreign protein can usually be recovered. Published reports on both types of expression system are summarized below.

Stable genomic transformation using genes encoding foreign antigens

Streptococcus mutans *spaA protein*

The first report of the concept of using a plant expression system for the production of an edible vaccine appeared in a patent application published under the International Patent Cooperation Treaty[10]. It described a means to express a surface protein (spaA) from *S. mutans* in tobacco plants to a level of approximately 0.02% of the total leaf protein; the gene had been stably inserted by *Agrobacterium*-mediated transformation. Data were presented on the oral immunogenicity of spaA produced in *Escherichia coli*, which stimulated the production of S-IgA in saliva. No further reports of these studies have been published.

Hepatitis B surface antigen (HBsAg)

The expression of HBsAg at levels equal to 0.01% of total soluble protein in tobacco has been demonstrated[1]. The tobacco-derived recombinant HBsAg (rHBsAg) was recovered from leaf extracts as a VLP with an average size of 22 nm, which is important because the particle form of HBsAg is required for immunogenicity[11]. The plant-derived VLPs mimic the appearance of recombinant yeast-derived HBsAg particles[12], which is the material that is used in the currently available recombinant vaccine for hepatitis B (Recombivax®; distributed by Merck, Sharpe, and Dohme). In addition, the plant-derived material had similar buoyant density and antigenicity to human- and yeast-derived HBsAg, indicating faithful preservation of protein folding characteristics in the plant system[1].

A crude extract of rHBsAg from plants was used in parenteral immunization studies with mice[13]. The extract caused an immune response that was similar to the one achieved with Recombivax®, and included all the IgG subclasses, as well as IgM. Because T-cell-mediated immunity is essential for the prevention of hepatitis B, the fidelity of the T-cell epitope expressed by the tobacco-derived rHBsAg was investigated[13]. T cells isolated from the lymph nodes of mice, which were primed by parenteral immunization with the tobacco-derived rHBsAg, could be stimulated to proliferate *in vitro* by the tobacco-derived and yeast-derived rHBsAg. In total, the studies of rHBsAg from plants conclusively demonstrate that B- and T-cell epitopes of HBsAg are preserved when the antigen is expressed in transgenic plants, and that the recombinant antigen is produced as a VLP that mimics the currently available commercial vaccine.

E. coli *heat-labile enterotoxin B subunit and cholera-toxin B subunit*

In developing countries, diarrheal disease is an important cause of mortality, especially among children. Bacteria which cause diarrhea include *Vibrio cholerae* and the related enterotoxigenic *E. coli*. An oral vaccine composed of the cholera-toxin B subunit (CT-B) with killed *V. cholerae* cells gives protection against cholera and enterotoxigenic *E. coli* (Ref. 14). However, international health organizations have not distributed this vaccine because the cost of production of CT-B is too high for developing countries to afford.

The heat-labile enterotoxin (LT) from *E. coli* is a multimeric protein that is structurally, functionally and antigenically very similar to cholera toxin (CT). X-ray crystallography[15] has been used to determine that LT has one A subunit (LT-A), with an M_r of 27 kDa, and a pentamer of B subunits (LT-B), each of which has an M_r of 11.6 kDa. Specific binding of the nontoxic LT-B pentamer to the G_{M1} gangliosides present on

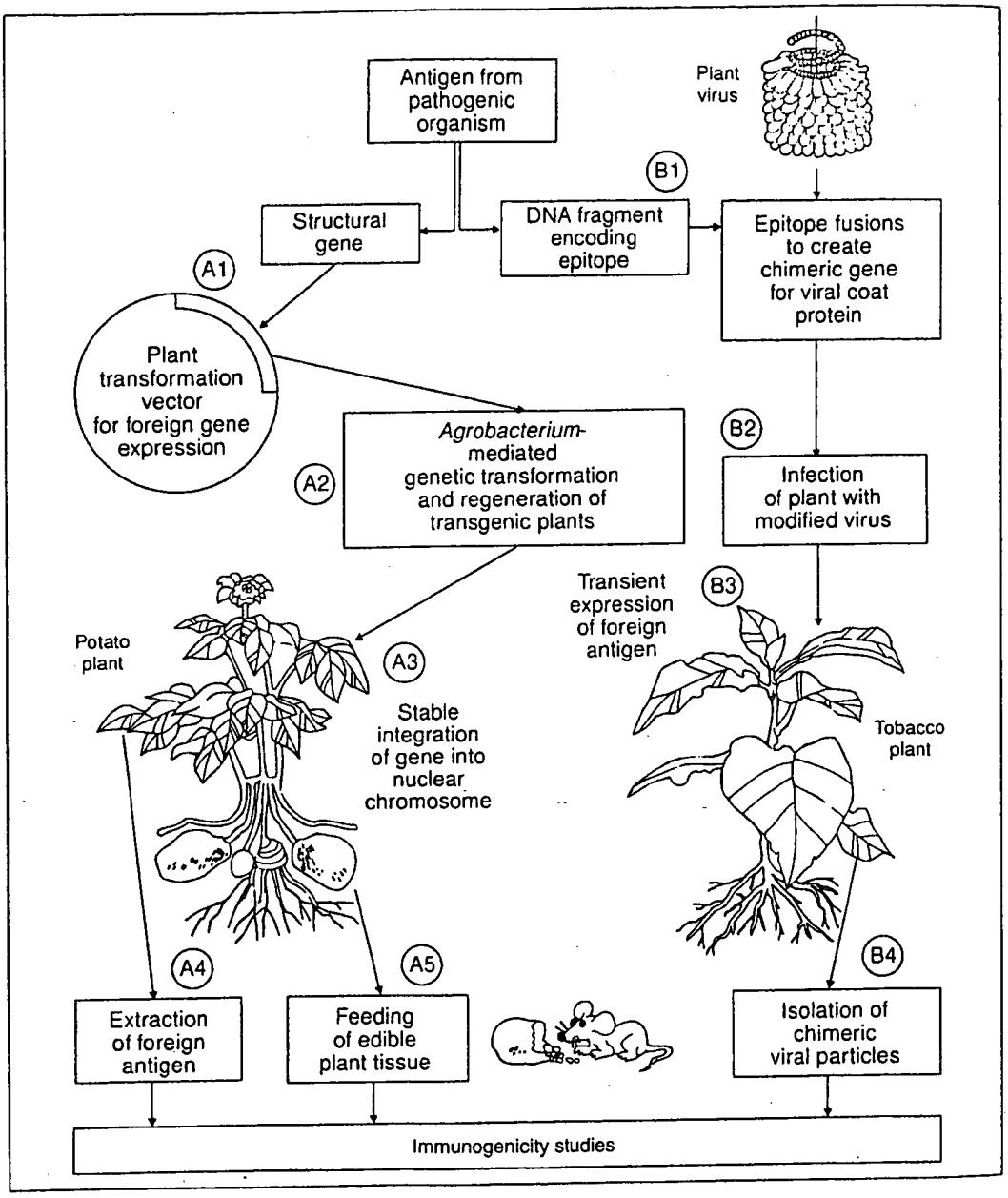


Figure 1

Strategies for the production of candidate vaccine antigens in plant tissues. Genes encoding antigens from pathogenic organisms (viruses, bacteria or parasites) that have been characterized and for which antibodies are available, can be handled in two ways. In one case, the entire structural gene is inserted into a plant transformation vector between 5′ and 3′ regulatory elements (**A1**); this will allow transcription and accumulation of the coding sequence in all, or selected, plant tissues. This vector is then used for the *Agrobacterium*-mediated transformation of plant cells (**A2**), or for stable integration of the expression cassette by other means, and regeneration of transgenic plants. The resulting plants contain the expression cassette stably integrated into the nuclear chromosomal DNA (**A3**), and can be used either for extraction and partial purification of the foreign antigen (**A4**), or for direct feeding of plant tissues (**A5**; in this case, a potato tuber) for assessment of immunogenicity. Alternatively, if epitopes within the antigen are identified, DNA fragments encoding these can be used (**B1**) to construct chimeric genes by fusion with a coat protein gene from a plant virus, e.g. tobacco mosaic virus (TMV) or cowpea mosaic virus (CPMV). The recombinant virus, or in the case of TMV and CPMV, even the viral RNA made *in vitro* from the plasmid clone, is then used to infect established plants (**B2**). Virus replication and systemic spread allow high-level transient expression of the chimeric coat protein in most plant tissues (**B3**). The viral particles, expressing the foreign epitope on their surfaces, are then purified (**B4**) and used for immunogenicity studies.

epithelial cell surfaces allows entry of the toxic LT-A subunit into cells. LT-B and CT-B are both potent oral immunogens[16].

LT-B has been expressed in transgenic plants, although the levels of expression were low[7]. The recombinant LT-B (rLT-B) produced in tobacco and potato showed enhanced accumulation when a C-terminal microsomal-retention sequence (Ser-Glu-Lys-Asp-Glu-Leu) was added. Sequestration of the rLT-B within microsomal vesicles may favor the association of individual subunits into the more stable pentameric form of LT-B. The tobacco-derived rLT-B appeared to be at least partially pentamerized, as judged by gel-permeation chromatography and its ability to bind gangliosides[7].

The oral immunogenicity of rLT-B was tested in mice and compared with bacterial LT-B (Ref. 7). When given orally to mice by gastric intubation, the plant-derived antigen stimulated humoral and mucosal immune responses, with titers comparable to the bacteria-derived LT-B. In addition, the antibodies produced against the tobacco-derived LT-B were able to neutralize LT activity, indicating the potential protective value of the immune response. The oral immunogenicity of unpurified rLT-B was also assessed by feeding raw transgenic potato tubers to mice. After only four feedings of 5 g tuber samples to mice, mucosal and serum antibodies were recovered. No immune response was observed in animals that were fed non-transformed tubers. This demonstrates that a food source containing a foreign antigen can induce oral immunization.

It should also be noted that CT and LT are excellent oral adjuvants, which stimulate immune responses against co-fed antigens[16,17]. Co-ordinate expression of A and B subunits to form the holotoxin in plants could enhance the vaccine value of other less immunogenic plant-expressed vaccine antigens expressed in the same tissues. This strategy is feasible because CT and LT function as adjuvants at concentrations well below those that cause diarrhea[16,17].

Transient expression of candidate vaccines using viral vectors

Using viral vectors for transient expression in plants represents a potentially useful means of producing high levels of recombinant antigens (Fig. 1). With tobacco mosaic virus (TMV), there are two ways to attain foreign-protein expression: (1) foreign-gene transcription, driven from a subgenomic promoter; and (2) fusion of foreign proteins or peptides with the capsid protein that normally coats the virus. The first strategy has been used to produce high-level expression of α-trichosanthin, an antiviral protein, in transfected plants[18], but the length of the foreign DNA insert that can be tolerated is undetermined. Capsid–protein fusions may be a better strategy, especially because the foreign protein is in particulate form (TMV virions), which is highly immunogenic[19]. Although the recombinant virus would need to be highly purified for parenteral administration, or partially purified for oral administration, this strategy may prove to be a cost-effective alternative to cell culture-based recombinant expression.

Malarial epitope fusions with TMV capsid protein

Turpen *et al.* have described a method for engineering the capsid protein of TMV as either internal or C-terminal fusions with peptides carrying epitopes derived from malarial sporozoites[19]. Both internal and C-terminal fusion constructs yielded high titers of genetically stable recombinant virus when used to infect tobacco plants. Antigenicity, measured by enzyme-linked immunosorbent assay (ELISA) and western blot, showed that the recombinant capsid proteins were recognized by the appropriate monoclonal anti-malarial antibodies.

Zona pellucida protein fusion with TMV capsid protein

The zona pellucida ZP3 protein of mammalian oocytes has been a target for immune contraception, and an epitope of 13 amino acids from murine ZP3 has now been expressed in plants as a fusion with TMV capsid protein[20]. The recombinant virus accumulated to high levels in infected plants, although systemic movement was somewhat slower than for the wild-type virus, and the viral particles were smaller. Mice immunized with the recombinant virus developed antibodies against ZP3 that were recruited to the zona pellucida; ovarian pathology was seen, but there was no observable effect on the fertility of the treated mice. Further work with extended epitopes or epitope mixtures may yield better results.

Cowpea mosaic virus capsid protein fusion

Cowpea mosaic virus (CPMV) has recently been developed as an expression system for the presentation of foreign peptides[21,22]. The advantages of CPMV are: high yield (1–2 g of virus per kg of host tissue), thermostability and ease of virus purification. CPMV is an icosahedral particle containing 60 copies each of large (L) and small (S) coat proteins. Earlier work showed that a foot and mouth disease virus epitope could be expressed as an S-protein fusion[21], but problems arose with the loss of the inserted RNA during serial passage. Modification of the chimeras resulted in genetically stable fusions, and the system has now been used to produce virus-expressing epitopes derived from human rhinovirus 14 and human immunodeficiency virus (HIV-1) that are immunogenic in test animals[22]. Furthermore, the antibodies raised against the CPMV–HIV chimera were able to neutralize three different strains of HIV-1 (Ref. 23). Inserts as long as 30 amino acids have been used in CPMV chimeras[21], although a particular sequence may have unpredictable effects on the systemic spread of the virus, thus compromising virus yield.

Future prospects

The demonstration that vaccine antigens can be produced in plants in their native, immunogenic forms opens exciting possibilities for the 'bio-pharming' of vaccines. If the antigens are orally active, food-based

'edible vaccines' could allow economical production and delivery in developing countries. Some engineering challenges remain, including maximizing the expression of the antigenic proteins, stabilizing the foreign protein during post-harvest storage in plant tissues, and enhancing the oral immunogenicity of some antigens.

Concerns about allergy and immune tolerance of orally applied antigens must be addressed as this technology develops, and will be resolved only by collaborative research efforts with medical specialists in these fields. It must be determined whether the levels of antigen required to induce the desired oral immune responses are less than the levels that could induce tolerance, as happens with proteins that occur as normal components of our food. If so, a delivery scheme must be developed to provide only the 'edible vaccine' as a medicinal product at the required dosage level and not as a routine food source. The specific oral response (immunogenicity versus tolerance) may be antigen specific and may, therefore, need to be considered on an individual vaccine basis. Research will be needed to determine which types of plants are most suitable for vaccine delivery; this must be coupled with the discovery of means to express the protein at the desired levels in the appropriate plant cells and tissues. Lastly, a thorough study of the relationship between immunogenic dose responses and antigen levels in various foods needs to be undertaken. In the short term, vaccines for animals are a more likely target for edible-vaccine technology than vaccines for humans, and studies in this area are likely to increase our understanding of the basic mechanisms, which can then be applied to the development of all vaccines.

References

1 Mason, H. S., Lam, D. M-K. and Arntzen, C. J. (1992) *Proc. Natl Acad. Sci. USA* 89, 11745–11749
2 Anonymous (1993) *The Children's Vaccine Initiative; Achieving the Vision* (Mitchell, V. S., Philipose, N. M. and Sanford, J. P., eds), National Academic Press
3 McGhee, J. R. and Kiyono, H. (1993) *Infect. Agents Disease* 2, 55–73
4 Castro, G. and Arntzen, C. J. (1993) *Am. J. Physiol.* 165, 599–610
5 de Aizpurua, H. J. and Russell-Jones, G. J. (1988) *J. Exp. Med.* 167, 440–451
6 Arntzen, C. J., Mason, H. S., Shi, J., Haq, T. A., Estes, M. K. and Clements, J. D. (1994) in *Vaccines '94: Modern Approaches to New Vaccines Including Prevention of AIDS* (Brown, F., Chanock, R. M., Ginsberg, H. S. and Lerner, R. A., eds), pp. 339–244, Cold Spring Harbor Laboratory Press
7 Haq, T. A., Mason, H. S., Clements, J. D. and Arntzen, C. J. (1995) *Science* 268, 714–716
8 Wenzler, H., Mignery, G., May, G. and Park, W. (1989) *Plant Sci.* 63, 79–85
9 May, G. D., Afza, R., Mason, H. S., Weicko, A., Novak, F. J. and Arntzen, C. J. (1995) *Bio/Technology* 13, 486–492
10 Curtiss, R. and Cardineau, G. A. (1990) *World Intellectual Property Organization*, PCT/US89/03799
11 Cabral, G. A. *et al.* (1978) *J. Gen. Virol.* 38, 339–350
12 Scolnick, E. M., McLean, A. A., West, D. J., McAleer, W. J., Miller, W. J. and Buynak, E. B. (1984) *J. Am. Med. Assoc.* 251, 2812–2815
13 Thanavala, Y., Yang, Y-F., Lyons, P., Mason, H. S. and Arntzen, C. J. (1995) *Proc. Natl Acad. Sci. USA* 92, 3358–3361
14 Clemens, J. D. *et al.* (1988) *J. Infect. Dis.* 158, 372–377
15 Sixma, T. K. *et al.* (1991) *Nature* 351, 371–377
16 Clements, J. D., Hartzog, N. M. and Lyon, F. L. (1988) *Vaccine* 6, 269–277
17 Holmgren, J., Lycke, N. and Czerkinsky, C. (1993) *Vaccine* 11, 1179–1184
18 Kumagai, M. H. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 427–430
19 Turpen, T. H., Reinl, S. J., Charoenvit, Y., Hoffman, S. L., Fallarme, V. and Grill, L. K. (1995) *Bio/Technology* 13, 53–57
20 Fitchen, J., Beachy, R. N. and Hein, M. B. *Vaccine* (in press)
21 Usha, R. *et al.* (1993) *Virology* 197, 366–374
22 Porta, C., Spall, V. E., Loveland, J., Johnson, J. E., Barker, P. J. and Lomonossoff, G. P. (1994) *Virology* 202, 949–955
23 McLain, L., Porta, C., Lomonossoff, G. P., Durrani, Z. and Dimmock, N. J. (1995) *AIDS Res. Hum. Retroviruses* 11, 327–334

If this isn't your copy of *Trends in Biotechnology* ...
or you would like additional copies of this Special Issue
Plant-product and crop biotechnology
... simply complete the order form below and send it to:

Elsevier Trends Journals, Oxford Fulfilment Centre, PO Box 800, Kidlington, Oxford, UK OX5 1GB. Tel: +44 (1865) 843300. Fax: +44 (1865) 843940.
OR Elsevier Trends Journals, 660 White Plains Road, Tarrytown, NY 10591-5153, USA. Tel: +1 (914) 524 9200. Fax: +1 (914) 333 2444.

Please send me ..... copies of **Plant-product and crop biotechnology**
Individuals: £16.00 + VAT* or $25.00 ..... Institutions: £36.00 + VAT* or $54.00 .....
Please enquire for bulk order discounts
Name .....
Address .....
.....
Zip/Postal code ..... Country .....
VAT registration No. .....
Cheque, Eurocheque or money order made payable to Elsevier
Credit card ☐ Visa/Barclaycard ☐ Mastercard/Access ☐ Am Express
Exp. date ..... Card No .....
Signature ..... Total payment (inc. VAT, except UK) .....
*EU non-UK customers only.
Non-VAT registered customers must pay VAT. Please add VAT (at your country's rate) to the price. VAT-registered customers please provide VAT registration no. EU countries and VAT equivalents: Belgium: BTW; Denmark: MOMS; France: TVA; Germany: MWST; Greece: FPA; Ireland: VAT; Luxembourg: TAV; Spain, IVA; Netherlands: BTW.

THIS PAGE BLANK (USPTO)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)